

FIGURE 2

MTDSPPPGHPEEKATPPGGTGHEGLSGGAADVASGVSGRHRARLPARPLGCVLSRAHGDVP
SESFVQRVYQPFLLTTCDGHRACSTYRTIYRTAYRRSPGLAPARPRYACCPGWKRTSGLPGAC
GAAICQPPCRNGSCVQPGRCRCPAGWRGDTQCSDVDECSARRGGCPQRCINTAGSYWCQCW

FIGURE 1

CCAGGTCCAAC TGCACCTCGGTTCTATCGATTGAATCCCCGGGGATCCTCTAGAGATCCCT
CGACCTCGACCCACGCGTCCGCCAAGCTGGCCCTGCACGGCTGCAAGGAGGCTCCTGTGGAC
CAGGCCAGGCAGGTGGGCCTCAGGAGGTGCTCCAGGCGGCCAGTGCGCCTGAGGCCCCAGC
AAGGGCTAGGGTCCATCTCCAGTCCCAGGACACAGCAGCGGCCACCATGGCCACGCTGGGC
TCCAGCAGCATCAGCAGCCCCCAGGACCGGGGAGGCACAGGTGGCCCCCACCACCCGGAGG
AGCAGCTCCTGCCCTGTCCGGGGGATGACTGATTCTCCTCCGCCAGGCCACCCAGAGGAGA
AGGCCACCCCGCCTGGAGGCACAGGCCATGAGGGGCTCTCAGGAGGTGCTGCTGATGTGGCT
TCTGTGTGTGGCAGTGGGCGGCACAGAGCACGCTTACCGCCCCGGCCGTAGGGTGTGTGCT
GTCCCGGGCTCACGGGGACCTGTCTCCGAGTCGTTCGTGCAGCGTGTGTACAGCCCTTCC
TCACCACCTGCGACGGGCACCGGGCCTGCAGCACCTACCGAACCATTTATAGGACCGCCTAC
CGCCCGAGCCCTGGGCTGGCCCCCTGCCAGGCCTCGCTACGCGTGCTGCCCCGGCTGGAAGAG
GACCAGCGGGCTTCTTGGGGCTGTGGAGCAGCAATATGCCAGCCGCGCATGCCGGAACGGAG
GGAGCTGTGTGCAGCCTGGCGCTGCCGCTGCCCTGCAGGATGGCGGGGTGACACTTGCCAG
TCAGATGTGGATGAATGCAGTGCTAGGAGGGGCGGCTGTCCCAGCGCTGCATCAACACGCG
CGGCAGTTACTGGTGCCAGTGTTGGGAGGGGCACAGCCTGTCTGCAGACGGTACACTCTGTG
TGCCCAAGGAGGGCCCCCAGGGTGGCCCCCAACCCGACAGGAGTGACAGTGCAATGAAG
GAAGAAGTGCAGAGGCTGCAGTCCAGGGTGGACCTGCTGGAGGAGAAGCTGCAGCTGGTGCT
GGCCCCACTGCACAGCCTGGCCTCGCAGGCACTGGAGCATGGGCTCCCGGACCCCGGCAGCC
TCCTGGTGCACTCCTTCAGCAGCTCGGCCGCATCGACTCCCTGAGCGAGCAGATTTCCTTC
CTGGAGGAGCAGCTGGGGTCTGTCTCTGCAAGAAAGACTCGTGACTGCCAGCGCCCCAGG
CTGGA CTGAGCCCTCACGCCGCCCTGCAGCCCCCATGCCCTGCCCAACATGCTGGGGGTC
CAGAAGCCACCTCGGGGTGACTGAGCGGAAGGCCAGGCAGGGCCCTCCTCCTTTTCTCCTC
CCCTTCCCTCGGGAGGGTCCCCAGACCCTGGCATGGGATGGGCTGGGATTTTTTTTGTGAAT
CCACCCCTGGCTACCCCCACCCTGGTTACCCCAACGGCATCCCAAGGCCAGGTGGGCCCTCA
GCTGAGGGAAGGTACGAGTTCCCTGCTGGAGCCTGGGACCCATGGCACAGGCCAGGCAGCC
CGGAGGCTGGGTGGGGCCTCAGTGGGGGCTGCTGCCTGACCCCCAGCACATAAAATGAAA
CGTGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGGGCGGCCGCGACTCT
AGAGTCGACCTGCAGAAGCTTGGCCGCCATGGCCCACTTGTTTTATTGCAGCTTATAATGGT
TACAAAT

FIGURE 3

CGCTCGCCCCGTCGCCCCCTCGCCTCCCCGAGAGTCCCCTCGCGGCAGCAGATGTGTGTGGG
GTCAGCCACGCGGGGACT**ATGGT**GAAATTCCCGCGCTCACGCACTACTGGCCCCGATC
CGGTTCTTGGTGCCCTGGGCATACCAACATAGCCATCGACTTCGGGGAGCAGGCCTTGAA
CCGGGGCATTGCTGCTGTCAAGGAGGATGCAGTGCAGATGCTGGCCAGCTACGGGCTGGCGT
ACTCCCTCATGAAGTTCTTACCGGTCCTATGAGTGACTTCAAAAATGTGGGCTGGTGTTT
GTGAACAGCAAGAGACAGGACAGGACCAAAGCCGTCTGTGTATGGTGGTGGCAGGGGCCATCGC
TGCCGTCTTTACACACTGATAGCTTATAGTGATTAGGATACTACATTATCAATAAACTGC
ACCATGTGGACGAGTCCGTGGGGAGCAAGACGAGAAGGCCCTTCTGTACTCTCGCCGCTTT
CCTTTTCATGGACCAATGGCATGGACCCATGCTGGCATTCTCTTAAAAACAAAATACAGTTT
CCTGGTGGGATGTGCCTCAATCTCAGATGTCTAGCTCAGGTGTTTTGTAGCCATTTTGC
TTCACAGTCACCTGGAATGCCGGGAGCCCCGTCTCATCCCGATCCTCTCCTTGTACATGGGC
GCATCTGTGCGCTGCACCACCCTGTGCCTGGGCTACTACAAGAACATTACGACATCATCCC
TGACAGAAGTGGCCCGAGCTGGGGGGAGATGCAACAATAAGAAAGATGCTGAGCTTCTGGT
GGCCTTTGGCTCTAATTCTGGCCACACAGAGAATCAGTCGGCCTATTGTCAACCTCTTTGTT
TCCCGGGACCTTGGTGGCAGTTCTGCAGCCACAGAGGCAGTGGCGATTTTGACAGCCACATA
CCCTGTGGGTACATGTCATACCGCTGGTGTGACGGAAATCCGTGCTGTGTATCCTGCTTTCG
ACAAGAATAACCCAGCAACAACCTGGTGAGCAGGACACACAGTCACGGCAGCCCATC
AAGAAGTTCACCTTCGTCTGCATGGCTCTGTCACTCAGCTCTGTTCTCGATGTTTTGGAG
TCTGTGTTGTTCTTTTGGGATCTTCTCCTTCTTCCAGTTCAGTCACAGTGAGGGCGCAT
CTCACCGGTTGGCTGATGACACTGAAGAAAACCTTCGTCTTGGCCCCAGCTCTGTGCTGCG
GATCATCTGTCCTCATCGCCAGCCTCGTGGTCTTACCCTACCTGGGGGTGACGGTGCAGACC
TGGGCGTGGGCTCCCTCCTGGCGGGCTTTGTGGGAGAATCCACCATGGTCGCCATCGCTGCG
TGCTATGTCTACCGGAAGCAGAAAAGAAGATGGAGAATGAGTCGGCCACGGAGGGGGAAGA
CTCTGCCATGACAGACATGCCTCCGACAGAGGAGGTGACAGACATCGTGGAAATGAGAGAGG
AGAATGAAT**TAA**GGCACGGGACGCCATGGGCACTGCAGGGACGGTCAGTCAGGATGACACTTC
GGCATCATCTCTCCCTCTCCCATCGTATTTTGTTCCCTTTTTTTGTTTTGTTTGGTAAT
GAAAGAGGCCTTGATTTAAAGGTTTCGTGTCAATTCTTAGCATACTGGGTATGCTCACACT
GACGGGGGGACCTAGTGAATGGTCTTACTGTGCTATGTAAAAACAAACGAAACAACCTGAC
TTCATACCCCTGCCTCACGAAAACCCAAAAGACACAGCTGCCTCACGGTTGACGTTGTGTCC
TCCCTCCCTGGACAACTCCTCTTGGAACCAAAGGACTGCAGCTGTGCCATCGCGCCTCGGT
ACCCCTGCACAGCAGGCCACAGACTCTCCTGTCCCCCTTCATCGCTCTTAAGAATCAACAGG
TTAAACTCGGCTTCTTTGATTTGCTTCCAGTCACATGGCCGTACAAAGAGATGGAGCCC
CGGTGGCCTCTTAATTTCCCTTCTGCCACGGAGTTCGAAACCATCTACTCCACACATGCAG
GAGCGGGTGGCACGCTGCAGCCCGAGTCCCCGTTACACATGGAACCGGACCTGTGAC
CACAGCAGGTGACAGATGGACAGAATCTCCCGTAGAAAGGTTTGGTTTGAATGCCCCGGG
GGCAGCAAATGACATGGTTGAATGATAGCATTTCACTCTGCGTTCTCCTAGATCTGAGCAA
GCTGTCACTTCTACCCCCACCGGTGATATATACATGAGCTAACTTTTTAAATTTGCAGATAA
GCGCATCTCCAGATTCAGACCCCTGCCGATGACTTTTCTGAAAGGCTTGCTTTTCCCTCGC
CTTTCTGAAGGTGCGATTAGAGCGAGTCACATGGAGCATCTAACTTTGCATTTTAGTTTTT
TACAGTGAATGAAGCTTTAAGTCTCATCCAGCATTTAATGCCAGGTGCTGTAGGGTAAC
TTTTGAAGTAGATATATTACTGGTTCTGCTATCCTTAGCTAACTCTGCGGTACAGGTAA
TTGAGAATGTACTACGGTACTTCCCTCCACACCATACGATAAGCAAGACATTTTATAACG
ATACAGAGTCACTATGTGGTCTCCTGAAATAACGCATTTCGAAATCCATGCACTGCAGTA
TATTTTCTAAGTTTTTGAAAGCAGGTTTTTCTTAAAAAAATTTAGACACGGGTTCACT
AAATTGATTTAGTCAGAATTCCTAGACTGAAAGAACCATAACAAAAAATTTTAAAGATA
TAAATATATGCTGTATATGTTATGTAATTTATTAGGCTATAATACATTTCTATTTTCG
ATTTTCAATAAAATGTCTCTAATACAAAAA

FIGURE 4

MVKFPALTHYWPLIRFLVPLGITNIAIDFGEQALNRGIAAVKEDAVEMLASYGLAYSLMKFF
TGPMSEDFKNVGLVFVNSKRDRTKAVLCMVVAGAAAVFHTLIAYSDLGYIINKLHHVDES
GSKTRRAFLYLAAFPFMDAMAWTHAGILLKHKYSFLVGCASISDVIAQVVFVAILLHSHLEC
REPLLIPILSLYMGALVRCCTLCLGYKNIHDIIPDRSGPELGGDATIRKMLSFWWPLALIL
ATQRI SRPIVNL FVSRDLGGSSAATEAVAILTATYPVGHMPYGWLTEIRAVYPAFDKNNPSN
KL VSTSN TVTAAHIKKFTFVCMALSLTLCFVMFWTPNVSEKILIDIIGVDFAFaelcVvPLR
IFSFFPVPTVRAHLTGWLMTLKKTFVLAPSSVLRIIVLIASLVLPYLGvHGATLGvGSLL
AGFVGESTMVAIAACYVYRKQKKMENESATEGEDSAMTDMPPTEEVTDIVEMREENE

Transmembrane domains:

amino acids 86-106, 163-179, 191-205, 237-253, 327-343, 357-374,
408-423, 431-445

FIGURE 5

CCTGACAGAAGTGCCCCGGAGCTGGGGGAGATNCAACATTAAGAAGATGCTGAGCTTCTGGT
GCCNTTTGGCTCTAATTCTGGCCACACAGAGAANCAGTCGGCCTATTGTCAACCTCTTTGTT
TCCCCGGACCTTGGTGGCAGTTCTGCAGCCACAGAGGCAGTGGCGATTTTGACAGCCACATA
CCCTGTGGGTACATGCCATACGGCTGGTTGACGGAAATCCGTGCTGTGTATCCTGCTTTTCG
ACAAGAATAACCCAGCAACAACTGGTGAGCACGAGCAACACAGTCACGGCGGCCACATC
AAGAAGTTCACCTTCGTCTGCATGGCTCTGTCACTCACGCTCTGTTTTCGTGATGTTTTGGAC
ACCCAACGTGTCTGNGAAAATCTTGATAGACATCATCGGAGTGGACTTTGCCCTTGCAGAAC
TCTGTGTGTTCTTTGCGGATCTTCTCCTTCTTCCAGTTCAGTCACAGTGAGGGCGCAT
CTCACCGGTGGCTGATGACACTGAAGAAAACCTTCGTC

FIGURE 6

TGACGGAATCCCGGGCTGGGTATCCTGGTTTNGACAAGATAAACCCCAAGCAANAAATTGGG
GAGCAGGGCAAAACAGTNACGGGCAGCCACATCAAGAAGTTCACCTTNGTTTGNATGGNTC
TGTCAACTCACGCTNTGTTTCGTGATGTTTTGGACACCCAAAGTGTGTGAGAAAATTTTGAT
AGACATNATCGGAGTGGANTTTGCCTTTGCAGAANTTTGNGNTGTTTCCTTTGCGGATTTTCT
CCTTTTCCCAGTTCAGTCACAGNGAGGGCGCATCTCACCGGNGGNTGATGACANTGAAG
AAAACCTTTGTCCTTGCCCCAGCTNTTTGGTGCGGATCATTGTCCTNATNGCCAGCCTTGT
GGTCCTACCCCTACCTGGGGGTGCACGGTGCGACCCCTGGGCGTGGGTTCCTCCTGGCGGGCA

09973403-101601

TATTTCCAGTTCCGGTCACGGGGAGGGCGCATNTACCGGGTGCGTGANGACACTGAAGAAA
ACCTTNGTCTCTTGCCCCAGNNTTGTGNTGCGGATNATCGTCTCATCGCCAGCCTNGTGGT
CCTACCCCTACCTGGGGGTGCACGGTGAGAC

FIGURE 8

GCCCCGCGCCCGCGCCGGGCGCCCCGAAGCCGGGAGCCACCGCCATGGGGGCGCTGCCTGGGA
GCCTGCTCCTCGTCTCAGCTGCGCGTCTCTGCCTCTGCGGCTCTGCCCCCTGCATCCTGTGCAG
CTGCTGCCCGCCAGCCGCAACTCCACCGTGAGCCGCCTCATCTTCAAGTTCTTCTCTTCC
TGGGGGTGCTGGTGTCCATCATTATGCTGAGCCCGGGCGTGAGAGTCTAGCTCTACAAGCTG
CCCTGGGTGTGTGAGGAGGGGGCCGGGATCCCCACCGTCTCGAGGGCCACATCGACTGTGG
CTCCTGCTTGGCTACCGCGCTGTCTACCGCATGTGCTTCGCCACGGCGGCCTTCTTCTTCT
TCTTTTTACCCCTGCTCATGCTCTGCGTGAGCAGCAGCCGGGACCCCGGGCTGCCATCCAG
AATGGGTTTTGGTTCTTTAAGTTCTGATCCTGGTGGGCCTCACCGTGGGTGCCTTCTACAT
CCCTGACGGCTCCTTACCAACATCTGGTTCTACTTCGGCGTCTGGGCTCCTTCTCTTCA
TCCTCATCCAGCTGGTGTCTGCTCATCGACTTTGCGCACTCCTGGAACCAGCGGTGGCTGGGC
AAGCCGAGGAGTGCATTCCCGTGCCTGGTACGAGGCCTTCTTCTTCACTCTCTCTT
CTACTGTCTGTGATCGCGGCGTGCGCTGATGTTTACTTACTGAGCCAGCGGCT
GCCACGAGGGCAAGGTCTTTCATCAGCCTCAACCTCACCTTCTGTGTCTGCGTGTCCATCGCT
GCTGTCTGCCCAAGGTCCAGGACGCCAGCCCAACTCGGGTCTGTGACGGCTCGGTCTAT
CACCTCTACACCATGTTTGTACCTGGTCAGCCCTATCCAGTATCCCTGAACAGAAATGCA
ACCCCATTTGCCAACCCAGCTGGGCAACGAGACAGTTGTGGCAGGCCCCGAGGGCTATGAG
ACCCAGTGGTGGGATGCCCCGAGCATTGTGGGCTCATCATCTTCTCTGTGCACCTCTT
CATCAGTCTGCGCTCCTCAGACCACCGCAGGTGAACAGCCTGATGCAGACCGAGGAGTCC
CACCTATGCTAGACGCCACACAGCAGCAGCAGCAGGTTGGCAGCCTGTGAGGGCCGGGCC
TTTGACAACGAGCAGGACGGCGTCACCTACAGCTACTCCTTCTTCCACTTCTGCCTGGTGT
GGCCTCACTGCACGTGATGATGACGCTACCAACTGGTACAAGCCCGGTGAGACCCGAAGA
TGATCAGCACGTGGACCGCCGTGTGGGTGAAGATCTGTGCCAGCTGGGACGGCTGCTCCTC
TACCTGTGGACCTGGTAGCCCCACTCCTCCTGCGCAACCGCGACTTCAGCTTAGGCAGCCT
CACAGCCTGCCATCTGGTGCCTCCTGCCACCTGGTGCCTCTCGGCTCGGTGACAGCCAACT
GCCCCCTCCCCACACCAATCAGCCAGGCTGAGCCCCACCCCTGCCCCAGCTCCAGGACCTG
CCCCTGAGCCGGGCTTCTAGTCTGATGCTTTCAGGCTCCGAGGAGCATCAGGCTCCTGCA
GAGCCCCATCCCCCGCCACACCCACAGGTGGAGTGCCTTCTCTTCCCCCTCCTCCTGT
TGCCATACTCAGCATCTCGGATGAAAGGGCTCCCTTGTCTCTCAGGCTCCACGGGAGCGGG
CTGCTGGAGAGAGCGGGAACTCCACCACAGTGGGCGATCCGGCACTGAAGCCCTGGTGT
CCTGGTCACTGCCCCAGGGGACCTGCCCCCTTCTGGACTTCGTGCCTTACTGAGTCTCT
AAGACTTTTTCTAATAACAAGCCAGTGCCTGTAAAAAAA

FIGURE 9

MGACLGACSLSCASCILCGSAPCILSCCPASRNSTVSRLIFTFFFLGLVLSIIMLSPGVE
SOLYKLPWVCEEAGIPTVLQGHIDCGSLLGYRAVYRMCFATAAFFFFFFTLLMLCVSSSRD
PRAAIQNGFWFFKFLILVGLTVGAFYIPDGSFTNIWFYFGVVGSLFLILQLVLLIDFAHSW
NQRWLKGAECDRAWYAGLFFFTLLFYLLSIAAVALMFMYYTEPSGCHEGKVFISLNLTFC
VCVSIAAVLPKVQDAQPNSGLLQASVITLYTMFVTWSALSSIPEQKCNPHLPTQLGNETVVA
GPEGYETQWWDAPSIVGLIIFLLCTLFISLRSSDHRQVNSLMQTEECPPMLDATQQQQQQVA
ACEGRAFDNEQDGVITYSYFFHFCLVLASLHVMMTLTNWYKPGETRKMISTWTAVVVKICAS
WAGLLLYLWTLVAPLLLLRNDRFS

Signal sequence:

amino acids 1-20

Transmembrane domains:

amino acids 40-58, 101-116, 134-150, 162-178, 206-223, 240-257,
272-283, 324-340, 391-406, 428-444

FIGURE 10

GAGCGAGGCCGGGGACTGAAGGTGTGGGTGTGAGCCCTCTGGCAGAGGGTTAACCTGGGTC
AAATGCACGGATTCTCACCTCGTACAGTTACGCTCTCCCGCGGCACGTCGCGAGGACTTGA
AGTCCTGAGCGCTCAAGTTTGTCCGTAGGTGAGAGAAAGGCCATGGAGGTGCCGCCACCGGC
ACCGCGGAGCTTTCTCTGTAGAGCATTGTGCCCTATTTCCCGAGTCTTTGCTGCCAAGCTG
TGACTGCCGATTCTGGAAGTCCTTGAGGAGCGTCAGAAGCGGCTTCCCTACGTCCCAGAGCCC
TATTACCCGGAATCTGGATGGGACCGCCTCCGGGAGCTGTTTGGCAAAGATGAACAGCAGAG
AATTTCAAAGGACCTTGCTAATATCTGTAAGACGGCAGCTACAGCAGGCATCATTGGCTGGG
TGTATGGGGGAATACCAGCTTTTATTTCATGCTAAACAACAATACATTGAGCAGAGCCAGGCA
GAAATTTATCATAACCGGTTTGATGCTGTGCAATCTGCACATCGTGCTGCCACACGAGGCTT
CATTGCTTATGGCTGGCGCTGGGGTTGGAGAACTGCAGTGTTTGTGACTATATTCAACACAG
TGAACACTAGTCTGAATGTATACCGAAATAAAGATGCCTTAAGCCATTTTGTAATTGCAGGA
GCTGTACAGGGAAGTCTTTTTAGGATAAACGTAGGCCTGCGTGGCCTGGTGGCTGGTGGCAT
AATTGGAGCCTTGCTGGGCACTCCTGTAGGAGGCTGCTGATGGCATTTCAGAAGTACGCTG
GTGAGACTGTTTCAGGAAAGAAAACAGAAGGATCGAAAGGCACTCCATGAGCTAAAACCTGGAA
GAGTGGAAGGCAGACTACAAGTTACTGAGCACCTCCCTGAGAAAATTGAAAGTAGTTTACG
GGAAGATGAACCTGAGAATGATGCTAAGAAAATTGAAGCACTGCTAAACCTTCCTAGAAACC
CTTCAGTAATAGATAAAACAAGACAAGGACTTGAAAGTGCTCTGAACCTGAAACTCACTGGAGA
GCTGAAGGGAGCTGCCATGTCCGATGAATGCCAACAGACAGGCCACTCTTTGGTCAGCCTGC
TGACAAATTTAAGTGCTGGTACCTGTGGTGGCAGTGGCTTGCTCTTGCTTTTTCTTTCTTT
TTTAACCTAAGAATGGGGCTGTTGTACTCTCACTTTACTTTATCCTTAAATTTAAATACATACT
TATGTTTGTATTAATCTATCAATATATGCATACATGGATATATCCACCACCTAGATTTTAA
GCAGTAAATAAAACATTTTCGCAAAAGATTAAAGTTGAATTTTACAGTTT

FIGURE 11

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA23318

><subunit 1 of 1, 285 aa, 1 stop

><MW: 32190, pI: 9.03, NX(S/T): 2

MEVPPPPAPRSFLCRALCLFPRVFAAEAVTADSEVLEERQKRLPYVPEPYYPESGWDRLRELF
GKDEQQRISKDLANICKTAATAGIIGWVYGGIPAFIHAKQQYIEQSQAIEYHNRFDAVQSAH
RAATRGFIRYGWRWGWRTAVFVTIFNTVNTSLNVYRNKDLSHFVIAGAVTGS�FRINVGLR
GLVAGGIIGALLGTPVGGLLMAFQKYAGETVQERKQKDRKALHELKLEEWKGRLQVTEHLPE
KIESSLREDEPENDAKKIEALLNLPRNPSVIDKQDKD

Important Features:

Signal Peptide:

amino acids 1-24

Transmembrane domains:

amino acids 76-96 and 171-195

N-glycosylation site:

amino acids 153-156

FIGURE 12

CGGAAGTCCCTTGAGGAGCGTCAGAAGCGGCTTCCCTACGTCCCAGAGCCCTATTACCCGGA
ATCTGGATGGGACCGCTCCGGGAGCTGTTTGGCAAAGATGAACAGCAGAGAATTTCAAAGGA
CCTTGCTAATATCTGTAAGACGGCAGCTACAGCAGGCATCATTGGCTGGGTGTATGGGGGAA
TACCAGCTTTTATTTCATGCTAAACAACAATACATTGAGCAGAGCCAGGCAGAAATTTATCAT
AACCGGTTTGATGCTGTGCAATCTGCACATCGTGCTGCCACACGAGGCTTCATTTCGTTTCATG
GCTGGCGCCGAACC

109707-204260

FIGURE 13

TCAAGTTTGTCCGTAGGTCGAGAGAAGGCCATGGAGGTGCCGCCACCGGCACCGCGGAGCTT
TTTTCTGTAGAGCATTGTGCCTATTTCCCCGAGTTTTTGTCTGCCGAAGCTGTGACTGCCGAT
TCGGAAGTCCTTGAGGAGCGTCAGAAGCGGCTTCCCTACGTCCCAGAGCCCTATTACCCGGA
ATTTGGATGGGACCGCCTCCGGGAGCTGTTTGGCAAAGATGAACAGCAGAGAATTTCAAAGG
ACCTTGCTGATATNTGTAAGACGGCAGCTACAGCAGGCATCATTGGCTGGGTGTATGGGGGA
ATACCAGCTTTTATTTCATGNTAAACAACAATACATTGAGCAGAGCCAGGCAGAAATTTATNA
TAACC

016501

[illegible]

FIGURE 15

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA39979

><subunit 1 of 1, 204 aa, 1 stop

><MW: 22147, pI: 8.37, NX(S/T): 3

MVCGGFACSKNCLCALNLLYTLVSLLLIGIAAWGIGFGLISSLRVVGVIAGIFLFLIALV
GLIGAVKHHQVLLFFYMIILLLVFIVQFSVSCACLALNQEQQGQLLEVGNWNTASARNDIQR
NLNCCGFRSVNPNDTCLASCVKSDHSCSPCAPIIGEYAGEVLRVFGGIGLFFSFTEILGVWL
TYRYRNQKDPNPSAFL

Signal Peptide:

amino acids 1-34

Transmembrane domains:

amino acids 47-63, 72-95 and 162-182

TGATTGGAGCTGTAAAAAANTCTTCAGGTGTTGTTNATTTTTTTTATATGATTATTCTGTGAANT
 TGTATTTATTGTTTCAGTTTTNTGTATCTTGCCTTGTTTAGCCNTGAACCAAGGAGCAACAGG
 GTCAGNTTNTGGAGGTTGGTTGGAAACAATACGGCAAGTGCTCGAAATGACATCCAGAGAAAT
 NTAAACTGCTGTGGGTTCCGAAGTGTTAACCCAAATGACACCTGTNTGGCTAGCTGTGTTAA
 AAGTGACCACTNGTGCTCGCCATGTGCTCCAATCATAGGAGAATATGCTGGAGAGGTTTTGA
 GATTTGTTGGTGGCATTGGCCTGTTNTTCAGTTTTACAGAGATCCTGGGTGTTTGGCTGACC
 TACAGATACAGGAACCAAG

FIGURE 17

AATCCCAAATCCCCAATTTTTTTGGNCTTTTTAGGGAAAGATGTGTTGTGGTAAAAAGTGT
TAGTATAAAAATGATAATTTACTTGTAGTCTTTTATGATTACACCAATGTATTCTAGAATAG
TTATGTCTTAGGAAATTGTGGTTTAATTTTTGACTTTTACAGGTAAGTGCAAAGGAGAAGTG
GTTTCATGAAATGTTCTAATGTATAATAACATTTACCTTCAGCCTCCCATCAGAATGGAACG
AGTTTTGAGTAATCCAGGAAGTATATCTATATGATCTTGATATTGTTTTATATAATTGAAG
TCTAAAAGACTGCATTTTTTAAACAAGTTAGTATTAATGCGTTGGCCACGTAGCAAAAAGAT
ATTTGATTATCTTAAAAATTGTTAAATACCGTTTTCATGAAAGTTCTCAGTATTGTAACAGC
AACTTGTCAAACCTAAGCATATTTGAATATGATCTCCATAATTTGAAATTGAAATCGTATT
GTGTGGAGGAAATGGCAATCTTATGTGTGCTGAAGGACACAGTAAGAGCACCAAGTTGTGCC
CCACTTGC

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 18

ATGATTATTCTGTTACTTGTATTTATTGTTTCAGTTTTATGGTATCTTGCGCTTGTTTAGCCC
CTGAAACCAGGAGCAACAGGGNNCAGCTTCCTGGAGGTTGGTTGGCAACAATCACGGCCAAG
TGACTCCGCAAATGACATCCCAGAGAAATCCTAAACTGCTGTGGGTTCCGAAGTGTTAACCC
AAATGACACCTGTCTGGCTNGCTGTGTTAAAAGTGACCACTCGTGCTCGCCATGTGCTCCAA
TCATAGGAGAATATGC

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 19

CAGTCACCATGAAGCTGGGCTGTGTCTCATGGCCTGGGCCCTCTACCTTTCCCTTGGTGTG
CTCTGGGTGGCCAGATGCTACTGGCTGCCAGTTTGTAGACGCTGCAGTGTGAGGGACCTGT
CTGCACTGAGGAGAGCAGCTGCCACACGGAGGATGACTTGA CTGATGCAAGGGAAGCTGGCT
TCCAGGTCAAGGCCTACACTTTCACTGAACCCCTCCACCTGATTGTGTCTATGACTGGCTG
ATCCTCCAAGGTCCAGCCAAGCCAGCTTTTGAAGGGGACCTGCTGGTTCTGCGCTGCCAGGC
CTGGCAAGACTGGCCACTGACTCAGGTGACCTTCTACCGAGATGGCTCAGCTCTGGGTCCCC
CCGGGCCCTAACAGGGAATTCTCCATCACCGTGGTACAAAGGCAGACAGCGGGCAGTACCAC
TGCAGTGGCATCTTCCAGAGCCCTGGTCTGGGATCCCAGAAACAGCATCTGTTGTGGCTAT
CACAGTCCAAGAAGCTGTTTCCAGCGCCAATTCTCAGAGCTGTACCCCTCAGCTGAACCCCAAG
CAGGAAGCCCCATGACCTGAGTTGTGAGCAAAAGTTGCCCTGCAGAGGTGAGCTGCCCGC
CTCCTCTTCTCTCTTCTACAGGATGGAAGGATAGTGCAAGCAGGGGGCTCTCCTCAGAATT
CCAGATCCCCACAGCTTCAGAAGATCACTCCGGGTCACTAGTGGTGTGAGGCAGCCACTGAGG
ACAACCAAGTTTGGAAACAGAGCCCCCAGCTAGAGATCAGAGTGCAGGGTGCTTCCAGCTCT
GCTGCACCTCCCACATTGAATCCAGCTCCTCAGAAATCAGCTGCTCCAGGAAGTGTCTCTGA
GGAGGCCCTGGGCCTCTGCCTCCGCCGCCAACCCCATCTTCTGAGGATCCAGGCTTTTCTT
CTCCTCTGGGGATGCCAGATCCTCATCTGTATCACCAGATGGGCCCTTCTTCTCAAACACATG
CAGGATGTGAGAGTCTCCTCGGTACCTGCTCATGGAGTTGAGGGAATTATCTGGCCACCA
GAAGCCTGGGACCACAAAGGCTACTGCTGAATAGAAAGTAAACAGTTTATCCATGATCTCACT
TAACCACCCCAATAAATCTGATTCTTTATTTTCTTTCCTGTCTGACATATGCATAAGTA
CTTTTACAAGTTGTCCAGTGTTTTGTGTTAGAATAATGTAGTTAGGTGAGTGTAATAAATTT
ATATAAAGTGAGAATTAGAGTTTAGCTATAATTGTGTATTCTCTCTTAACACAACAGAATTC
TGCTGTCTAGATCAGGAATTTCTATCTGTTATATCGACCAGAATGTTGTGATTTAAAGAGAA
CTAATGGAAGTGGATTGAATACAGCAGTCTCAACTGGGGGCAATTTTGCCCCCAGAGGACA
TTGGGCAATGTTTGGAGACATTTTGGTCATTATACTTGGGGGGTTGGGGGATGGTGGGATGT
GTGTCTACTGGCATCCAGTAAATAGAAGCCAGGGGTGCCGTAAACATCCTATAATGCACAG
GGCAGTACCCCAACAGAAAAATAATCTGGCCCAAATGTGAGTTGTA CTGAGTTTGAGAAA
CCCCAGCCTAATGAAACCTAGGTGTTGGGCTCTGGAATGGGACTTTGTCCCTTCTAATTAT
TATCTCTTTCCAGCCTCATTCAGCTATTCTTACTGACATACCAGTCTTTAGCTGGTGTATG
GTCTGTTCTTTAGTTCTAGTTTGATCCCCCTCAAAGCCATTATGTTGAAATCCTAATCCCC
AAGGTGATGGCATTAAAGAAGTGGGCCTTTGGGAAGTGATTAGATCAGGAGTGCAGAGCCCTC
ATGATTAGGATTAGTGCCCTTATTTAAAAAGGCCCCAGAGAGCTAACTCACCTTCCACCAT
ATGAGGACGTGGCAAGAAGATGACATGTATGAGAACCAAAAACAGCTGTGCGCAAAACACCG
ACTCTGTCGTTGCCCTTGATCTTGAACCTCCAGCCTCCAGAACTATGAGAAATAAAATTCTGG
TTGTTTGTAGCCTAA

U3976407-10344

Parameter	Value	Unit
Temperature	25	°C
Pressure	1.0	atm
Flow rate	1.0	L/min
Concentration	0.1	mol/L
Time	10	min
Volume	10	L
Mass	10	g
Energy	10	J
Power	10	W
Frequency	10	Hz
Wavelength	10	nm
Angle	10	°
Distance	10	m
Area	10	m ²
Volume	10	m ³
Mass	10	kg
Energy	10	J
Power	10	W
Frequency	10	Hz
Wavelength	10	nm
Angle	10	°
Distance	10	m
Area	10	m ²
Volume	10	m ³
Mass	10	kg
Energy	10	J
Power	10	W
Frequency	10	Hz
Wavelength	10	nm
Angle	10	°
Distance	10	m
Area	10	m ²
Volume	10	m ³
Mass	10	kg
Energy	10	J
Power	10	W
Frequency	10	Hz
Wavelength	10	nm
Angle	10	°
Distance	10	m
Area	10	m ²
Volume	10	m ³
Mass	10	kg
Energy	10	J
Power	10	W
Frequency	10	Hz
Wavelength	10	nm
Angle	10	°
Distance	10	m
Area	10	m ²
Volume	10	m ³
Mass	10	kg
Energy	10	J
Power	10	W
Frequency	10	Hz
Wavelength	10	nm
Angle	10	°
Distance	10	m
Area	10	m ²
Volume	10	m ³
Mass	10	kg
Energy	10	J
Power	10	W
Frequency	10	Hz
Wavelength	10	nm
Angle	10	°
Distance	10	m
Area	10	m ²
Volume	10	m ³
Mass	10	kg
Energy	10	J
Power	10	W
Frequency	10	Hz
Wavelength	10	nm
Angle	10	°
Distance	10	m
Area	10	m ²
Volume	10	m ³
Mass	10	kg
Energy	10	J
Power	10	W
Frequency	10	Hz
Wavelength	10	nm
Angle	10	°
Distance	10	m
Area	10	m ²
Volume	10	m ³
Mass	10	kg
Energy	10	J
Power	10	W
Frequency	10	Hz
Wavelength	10	nm
Angle	10	°
Distance	10	m
Area	10	m ²
Volume	10	m ³
Mass	10	kg
Energy	10	J
Power	10	W
Frequency	10	Hz
Wavelength	10	nm
Angle	10	°
Distance	10	m
Area	10	m ²
Volume	10	m ³
Mass	10	kg
Energy	10	J
Power	10	W
Frequency	10	Hz
Wavelength	10	nm
Angle	10	°
Distance	10	m
Area	10	m ²
Volume	10	m ³
Mass	10	kg
Energy	10	J
Power	10	W
Frequency	10	Hz
Wavelength	10	nm
Angle	10	°
Distance	10	m
Area	10	m ²
Volume	10	m ³
Mass	10	kg
Energy	10	J
Power	10	W
Frequency	10	Hz
Wavelength	10	nm
Angle	10	°
Distance	10	m
Area	10	m ²
Volume	10	m ³
Mass	10	kg
Energy	10	J
Power	10	W

><MW: 38899, pI: 5.21, NX(S/T): 0

amino acids 1-17

amino acids 12-33

amino acids 353-355

FIGURE 21

CCCACGCGTCCGCCCCACGCGTCCGCCCCACGGGTCCGCCCCACGCGTCCGGGCCACCAGAAGTT
TGAGCCTCTTTGGTAGCAGGAGGCTGGAAGAAAGGACAGAAGTAGCTCTGGCTGTGATGGGG
ATCTTACTGGGCTGCTACTCTGGGGCACCTAACAGTGGACACTTATGGCCGTCCCATCCT
GGAAGTGCACAGAGTGTAAACAGGACCTTGGAAAGGGGATGTGAATCTTCCCTGCACCTATG
ACCCCTGCAAGGCTACACCCAAGTCTTGGTGAAGTGGCTGGTACAACGTGGCTCAGACCCCT
GTCACCATCTTTCTACGTGACTCTTCTGGAGACCATATCCAGCAGGCAAAGTACCAGGGCCG
CCTGCATGTGAGGCCACAAGGTTCCAGGAGATGTATCCCTCCAATTGAGCACCCCTGGAGATGG
ATGACCGGAGCCACTACACGTGTGAAGTCACTGGCAGACTCCTGATGGCAACCAAGTCGT
AGAGATAAGATTACTGAGCTCCGTGTCCAGAAACTCTCTGTCTCCAAGCCCACAGTGACAAC
TGGCAGCGGTTATGGCTTCAAGGTGCCCCAGGGAATGAGGATTAGCCTTCAATGCCAGGCTC
GGGGTTCTCCTCCCATCAGTTATATTTGGTATAAGCAACAGACTAATAACCAGGAACCCATC
AAAGTAGCAACCCCTAAGTACCTTACTCTTCAAGCCTGCGGTGATAGCCGACTCAGGCTCCTA
TTTCTGCACTGCCAAGGGCCAGGTTGGCTCTGAGCAGCACAGCGACATTGTGAAGTTTGTGG
TCAAAGACTCCTCAAAGCTACTCAAGACCAAGACTGAGGCACCTACAACCATGACATACCCC
TTGAAAGCAACATCTACAGTGAAGCAGTCTTGGGACTGGACCACTGACATGGATGGCTACCT
TGGAGAGACCAGTGCTGGGCCAGGAAAGAGCCTGCCTGTCTTTGCCATCATCCTCATCATCT
CCTTGTGTGTATGGTGGTTTTTACCATGGCCTATATCATGCTCTGTGGAAGACATCCCAA
CAAGAGCATGTCTACGAAGCAGCCAGGTAAGAAAGTCTCTCCTCTTCCATTTTGACCCCGT
CCCTGCCCTCAATTTTGATTACTGGCAGGAAATGTGGAGGAAGGGGGGTGTGGCACAGACCC
AATCCTAAGGCCCGAGGCCCTTCAAGGTCAGGACATAGCTGCCTTCCCTCTCTCAGGCACCTT
CTGAGGTTGTTTTGGCCCTCTGAACACAAAGGATAAATTTAGATCCATCTGCCTTCTGCTTCC
AGAATCCCTGGTGGTAGGATCCTGATAATTAATTGGCAAGAATTGAGGCAGAAGGGTGGGA
AACCAGGACCACAGCCCCAAGTCCCTTCTTATGGGTGGTGGGCTCTTGGGCCATAGGGCACA
TGCCAGAGAGGCCAACGACTCTGGAGAAACCATGAGGGTGGCCATCTTCGCAAGTGGCTGCT
CCAGTGATGAGCCAACCTTCCAGAATCTGGGCAACAACACTCTGATGAGCCCTGCATAGGA
CAGGAGTACCAGATCATCGCCAGATCAATGGCAACTACGCCCGCTGCTGGACACAGTTCC
TCTGGATTATGAGTTTCTGGCCACTGAGGGCAAAAGTGTCTGTTAAAAATGCCCCATTAGGC
CAGGATCTGCTGACATAATTGCCTAGTCAGTCTTGCCTTCTGCATGGCCTTCTTCCCTGCT
ACCTCTCTTCTGGATAGCCCAAAGTGTCCGCCTACCAACACTGGAGCCGCTGGGAGTCACT
GGCTTTGCCCTGGAATTTGCCAGATGCATCTCAAGTAAGCCAGCTGCTGGATTTGGCTCTGG
GCCCTTCTAGTATCTCTGCCGGGGCTTCTGGTACTCCTCTCTAAATACCAGAGGGAAGATG
CCCATAGCACTAGGACTTGGTCATCATGCCCTACAGACACTATTCAACTTTGGCATCTTGCCA
CCAGAAGACCCGAGGGAGGCTCAGCTCTGCCAGCTCAGAGGACCAGCTATATCCAGGATCAT
TTCTCTTTCTTCAAGGCCAGACAGCTTTTAATTGAAATTGTATTTCACAGGCCAGGGTTCA
GTTCTGCTCCTCCACTATAAGTCTAATGTTCTGACTCTCTCCTGGTGCTCAATAAATATCTA
ATCATAACAGC

```
><subunit 1 of 1, 321 aa, 1 stop
```

MGILLGLLLLLGHLTVDTYGRPILEVPEPSVTGPKWGDVNLCTYDPLQGYTQVLVKWLVRGS
DPVTFIFLRDSSGDHIQQAQYQGRHLVSHKVPDVSLLSTLEMDDRSHYTCVETWQTDPDGNQ
VVRDKITELRVQKLSVSKPTVTGSGYGFTVPQGMRLISLQCQARGSPPISYIWKQOTNNQE
PIKVATLSTLLFKPAVIADSGSYFCTAKGVQVGEQHSIDIVKFVVKDSSKLLKTKTEAPTMT
YPLKATSTVKQSDWTTMDGYLGSETSAGPGKSLPVFAILLIISLCCMVVFTMAYIMLCRKT
SOEHVYEAAAR

amino acids 1-19

amino acids 149-152

amino acids 282-300

FIGURE 23

GC CGCCGGGAGCCCATCTGCCCCAGGGGCACGGGGCGGGGCCGGCTCCCGCCCGGCACAT
GGCTGCAGCCACCTCGCGCGCACCCCGAGGCGCGCGCCAGCTCGCCCGAGGTCGGTCCGGA
GGCGCCCGGCCGCCCCGGAGCCAAGCAGCAACTGAGCGGGGAAGCGCCCGCTCCGGGGATC
GGGATGTCCTCTCCTTCTCCTCTTGTCTAGTTTCTACTATGTTGGAACCTTGGGGACTCA
CACTGAGATCAAGAGAGTGGCAGAGGAAAAGGTCACTTTGCCCTGCCACCATCAACTGGGGC
TTCCAGAAAAAGACACTCTGGATATTGAATGGCTGCTCACCGATAATGAAGGGAACCAAAA
GTGGTGATCACTTACTCCAGTCGTCTGTCTACAATAACTTGACTGAGGAACAGAAGGGCCG
AGTGGCCTTTGCTTCCCAATTTCTGGCAGGAGATGCCTCCTTGCAATTGAACCTCTGAAGC
CCAGTGATGAGGGCCGGTACACCTGTAAGGTTAAGAATTGAGGGCGCTACGTGTGGAGCCAT
GTCATCTTAAAGTCTTAGTGAGACCATCCAAGCCCAAGTGTGAGTTGGAAGGAGAGCTGAC
AGAAGGAAGTGACCTGACTTTGAGTGTGAGTCATCCTCTGGCACAGAGCCCATTTGTGTATT
ACTGGCAGCGAATCCGAGAGAAAAGGGAGAGGATGAACGTCTGCCCTCCCAATCTAGGATT
GACTACAACCACCCTGGACGAGTTCTGCTGCAGAATCTTACCATGTCTCTACTCTGGACTGTA
CCAGTGCACAGCAGGCAACGAAGCTGGGAAGGAAAGCTGTGTGGTGCGAGTAACCTGTACAGT
ATGTACAAAGCATCGGCATGGTTGCGAGGAGCAGTGACAGGCATAGTGCTGGAGCCCTGCTG
ATTTTCTCTTGGTGTGGCTGCTAATCCGAAGGAAAGACAAAGAAAGATATGAGGAAGAAGA
GAGACCTAATGAAATTCGAGAAGATGCTGAAGCTCCAAAAGCCCGTCTTGTGAAACCCAGCT
CCTCTTCTCAGGCTCTCGGAGCTCACGCTCTGGTTCTTCTCCACTCGCTCCACAGCAAAT
AGTGCCTCACGCAGCCAGCGGACACTGTCAACTGACGCAGCACCCAGCCAGGGCTGGCCAC
CCAGGCATACAGCCTAGTGGGGCCAGAGGTGAGAGGTTCTGAACCAAAGAAAGTCCACCATG
CTAATCTGACCAAAGCAGAAACACACCCAGCATGATCCCAGCCAGAGCAGAGCCTTCCAA
ACGGTCTGAAATTACAATGGACTTGACTCCCACGCTTTCTAGGAGTCAGGGTCTTTGGACTC
TTCTCGTCATTGGAGCTCAAGTCACCAGCCACACAACCAGATGAGAGGTCATCTAAGTAGCA
GTGAGCATTGCACGGAACAGATTAGATGAGCATTCTTCTTATACAATACCAACAAGCAAA
AGGATGTAAGCTGATTCATCTGTAAAAAGGCATCTTATTGTGCCCTTAGACCAAGATAAGGG
AAAGCAGGAGTCCAAATCTATTGTGACCAGGACCTGTGGTGAGAAGGTTGGGGAAAGGTG
AGGTGAATATACCTAAAACCTTTAATGTGGGATATTTGTATCAGTGCTTTGATTACAATT
TTCAAGAGGAAATGGGATGCTGTTTGTAAATTTCTATGCAATTTCTGCAAACTTATTGGATT
ATTAGTTATTAGACAGTCAAGCAGAACCCACAGCCTTATTACACCTGTCTACACCATGTAC
TGAGCTAACCACTTCTAAGAAACTCCAAAAAGGAAACATGTGTCTTCTATTCTGACTTAAC
TTCATTTGTATAGGTTTGGATATTAAATTTCAAGGGGAGTTGAAATAGTGGAGATGGAGA
AGAGTGAATGAGTTTCTCCACTCTATACTAATCTCACTATTTGTATTGAGCCCAAAATAAC
TATGAAAGGAGACAAAAATTTGTGACAAAGGATGTGAAGAGCTTCCATCTTCATGATGTT
ATGAGGATTGTTGACAAACATTAGAAATATATAATGGAGCAATTGTGGATTTCCCCTCAAT
CAGATGCCTCTAAGGACTTTCTCTGCTAGATATTTCTGGAAGGAGAAAAATACAACATGTCTATT
TATCAACGTCCTTAGAAAGAATCTTCTAGAGAAAAAGGGATCTAGGAATGCTGAAAGATTA
CCCAACATACCATTATAGTCTCTTCTTCTGAGAAAAATGTGAAACCAGAATTGCAAGACTGG
GTGGACTAGAAAGGAGATTAGATCAGTTTTCTCTTAATATGTCAAGGAAGGTAGCCGGGCA
TGGTGCCAGGCACCTGTAGGAAAAATCCAGCAGGTGGAGGTTGCAGTGAGCCGAGATTATGCC
ATTGCACTCCAGCCTGGGTGACAGAGCGGGACTCCGTCTC

CCGTTTCTTGGTGTGGCTGCTAATCCGAAGGAAAGACAAAGAAAGATATGAGGAAGAAGA

TGTGGTTCCTTTGCTCTCTCTCGCCAGCAGTCTCCTCCCTGGTTTCTCTCTACGCGCTGTGCGGAGGAGAGACCCCGGA
 GACGCGGGCTGTCAGTCGCGCGGCGTCTTCCCGCGCTGCGCGGCGCTCGCGCTGGGACAGGTGTCAGCGCCCTTAG
 AGCCCTCCCTTTCGCGCTCCCTCTCTCTCGCCGCGCGGACAGTCACATGGGGTTGTGAGAGTAGATGGGCTCCCG
 GCGCGGGAGCGGCGCGGTGGATGTCGGCGCTGGGACGAAGACCGCCGCAATTCACGTCCTCCCGCGCGCCCGCGCG
 CCCCCTGCAGTTCCCGGTTTTCAGCTATGGGACCTCTCGAGCAGGACGACCGCCCTCCGCTCTCTGACGCGCAT
 GCCCGCCGAGCCACGACCCAGATGATCGGGGCTCCCTTCTCTGCTTGGATTCCTTAGCACCCACACAGCTCAG
 CCGAAGACGAAAGGCCCTGGAATCTCATTTGGCACATACCCGCTGTGACCGTCCACCGCGGCAGGTGTAACTGT
 GACAAGTGTTCAGAGGGAACCTATGTCTCTTGAGCATTTGACCAACAAAGCCTGGCGCTCTGACAGAGTTGCC
 GTGGGACACTTTACAGGCATAGAAATGGCATAGGAGAAGATGCCATGACTAGTGTACGATCCAGCTCCCATGGCCAACT
 ATTGAGAATAATCTTTGTCTGTCTGTGATGACCGAAGATCCATGTCTGCCACTGGGATGTTTCAGTCTAAAGCT
 ACTGTGTGCCCTGACATGGTGTCTCTGTGGTGTGGGGTGTGGCGGAAGAGGACAGAGACTAGAGATGTGGGG
 GTGAAGCAGTGTCTCGGGGTACCTCTCAGATGTGCTCTAGTGTGATGAAGTACAAAGCATACACAGACTGT
 CTGAGTCAGAACCTGGTGGTGATCAAGCCGGGGACCAAGGAGACAGACAACGTCGTGTGGCACACTCCCGTCCCTC
 TCAGACTCCACCTCACTTCCCTCGGACAGCCACTTCTTCAGCGCCCTGAGACAGATGGAACCCATGAAGTCCCT
 TCTCCACTTATGTGTTCCAAAGGACTGAATCAAGAGATCCAATCTCTGCTCTGTGAGCAAAAGGTACTGT
 AGTAGAGATCCGGAAGGAGACAGCTCCCTGCACACAAAGCTCAGCAGGGGGGAAGGAAGCTGAAACAGACACTGC
 CCAAACCTTCAGGTATCAACCAACGACAGCGCCCAACAGACACATCTGAAAGTGTGTGCTGCTCATGAG
 GCAACTGTGGGGCGAGAACTCCAGCACGCCCTCAAGSGCCCAAGGGGGACATCTAGACAGACACTTACAAG
 CATTTTGAACATTAATGAGCATTTGCCCTGGATGTGTGCTTTCTGTCTGTGTGTGTGTGTGTGTGTGTGTGT
 TGCAGTATCCGGAAGACTCGAGGACTCTGAAAGAGGGGGCCCGCGAGGATCCAGTGCATTGTGGAAGAGCA
 GGGCTGAAGAAATCATGATCTCAACACCGGAACCGGGAGGAAGATGGATCTACTACTGCAATGCCCATGGTATCGAT
 ATCTCAAGCTGTGTAGACGCGCAAGTGGGAAGCAGCTGGAAAGATATCTACGTTTCTTTGCAATGCGAGTGA
 AGGGAGAGTGTGTGCTTTCTCCAAATGGGACACAGCCGACACAGCGCGGCTACGCGACTCTGAGCATCTGGAC
 ATCCGGGGCCCGAGGCGCAGCTCGCCAGCATATAGCGCCTCGCGCAGCACCGGAGAAACGATTTGTGTGAG
 AAGATTCTGTGGGCTGATGGAAGACACCACCCAGCTGGAAACTGAAAACTAGTCTCTCCGATGAGCCCCAGCCGC
 CTTAGCCCGAGGCCCATCCCGAGCCCAAGCGGAACTTGAAATTCGCTCTCTCTGACGTTGGAGCTCCCCCA
 CAGGACAGAAAGCAAGGGCTCTCTCGTGATGAGTGGAGCGCCCTCTCGCTGTGATCATCTACATCAGCGGCTCC
 TCGCGCTGAGCAGGAACGGTTCCTTTATTAACAAGAAAGGAAGACAGAGTGTGGCGAGGTACGCTCGTGCAC
 CCTGTGATCTGAGGACCTATTTTGATGACATGCTCCACTTTCTAAATCTGAGGAGTGTGGGGTGTATGAAGAG
 ATTCCCGAGGCTGAGGACAAACTAGACCCGCTATTCGAAATTAATTGAGATCAAGAGCAGGAAGCCAGGCACAGCC
 CTCTCGACTCTGTTTATAGGCATCTTCTGTACCTGCTGTGATAGACATAGGGGATCTGCATTTGGAATTTACTA
 ATTTAGTGCAGGGGTGTTTTTTAAATTTTCTCTGTGTTGATTTTGTGTTTGGGGTGTGTGTGTGTGTTTGT
 GTGTGTGTGTGTGTGTGTGTGTGTGTGTTTAAACAGAGAATATGGCCAGTGTCTGAGTTCTTCTCTCTCT
 TCTCTCTCTTTTTTTAAATAACTCTCTGGGAAGTGTGTTTAAAGCCTTTGCGAGGTGTAATTTGTGA
 ATACCCACCACTAAAGTTTTTAAGTTCATATTTCTCCATTTTGGCTCTTATGATTTTTCAGATTAATCTG
 TGACCTTTAAATTTACTTAACCTACCTAATTCAGTGTGACTTTGCCACACATGATTGTGAGGCTCTTAAC
 TTCTTAAAGTATAATGGCATCTTGTGAATCTATAAGCAGTCTTATGTCTTAAACATTACACACTATCTTTT
 AAAAAACAATATTACTATTTTATATGTTTGTCTTTATAAATTTCTTAAAGATTAAAGAAATTTAAGA
 CCCCATTGAGTACTGTATAGCAATCAACTTTGAGTATCTTTTAAATATGCTCTGTAGTATGATATTAATTCAG
 CTGAAACTTGACACAATATGCTGTGATGTGTGTTTACCCTGGACACCGTGAGTAAGTGTGATTGTTGTATC
 TCTCTTATGTCAATAATGCTCTGGGTGTGGAAGAAATGAATCTCAAGCCATCAGGATTTGCTATTTAAGTGGCT
 GACCACTGGGCGCACCAAGGACTTGAACCTCACCTTTTAGGATTTAGCTGTCTGGAACACATGCTGTCACCTT
 GGAAGTCAAAATCAAGTGCAGCTGGCGCCCTTCCATAGAGAAATTTGCCAGCTTTCGTTTAAAGATGTCTTG
 TTTTTATATACACATAATCAATAGGTCCAATCTGCTCTCAAGGCTTGGTCTCGTGGGATTCTTCCACCAATT
 ACTTTAATTAATAATGGCTGCAACTGTGAAGAACCCCTGTGTGATATTTTCCAACTATGCTCCCATTTACAAGAT
 TACCTTGACCTGCTAGTTCGCACTTCCAAATGCAAGGTGGCGTGCATCTTTGTGTGGGTGGGGTTTGTGG
 GTAGTGTGGAAGGACCGATATCAGAAAAATGCCTTCAGTGTACTAATTTATTAATTAACATTAGGTGTGTGTA
 AAAAAAAA

FIGURE 27

ATGGGAAGCCAGTAACACTGTGGCCTACTATCTCTTCCGTGGTGCCATCTACATTTTTGGGA
CTCGGGAATTATGAGGTAGAGGTGGAGGCGGAGCCGATGTCAGAGSTCCTGAAATAGTCAC
CATGGGGGAAAATGATCCGCTGCTGTTGAAGCCCCTTCTCATTCCGATCGCTTTTTGGCC
TTGATGATTTGAAAATAAGTCTGTGACACAGATGCAGATGCTGTTGCTGCACAGATCCTG
TCACTGCTGCCATTGAAGTTTTTTCCAATCATCGTCATTGGGATCATTGCAATTGATATTAGC
ACTGGCCATTGGTCTGGGCATCCACTTCGACTGCTCAGGGAAGTACAGATGTCGCTCATCCT
TTAAGTGATCGAGCTGATAGCTCGATGTGACGGAGTCTCGGATTGCAAAAGACGGGGAGGAC
GAGTACCGCTGTGTCCGGGTGGGTGCTCAGAATGCCGTGCTCCAGGTGTTACAGCTGCTTC
GTGGAAGACCATGTGCTCCGATGACTGGAAGGGTCACTACGCAAAATGTTGCCCTGTGCCAAC
TGGGTTTTCCCAAGCTATGTGAGTTCAGATAACCTCAGAGTGAGCTCGCTGGAGGGGAGTTT
CGGGAGGAGTTTTGTGTCCATCGATCACCTCTTGCCAGATGACAAGGTGACTGCATTACACCA
CTCAGTATATGTGAGGGAGGGATGTGCCCTCTGGCCACGTGGTTACCTTGCAGTGCACAGCCT
GTGGTCATAGAAGGGGCTACAGCTCACGCATCGTGGGTGGAAACATGTCCTTGCTCTCGCAG
TGGCCCTGGCAGGGCAGCCTTCAGTTCCAGGGCTACCACCTGTGCGGGGGCTCTGTATCAC
GCCCTGTGGATCATCACTGCTGCACACTGTGTTTATGACTTGTACCTCCCAAGTCATGGA
CCATCCAGGTGGGTCTAGTTTCCCTGTTGGACAATCCAGCCCATCCCACTTGGTGGAGAAG
ATTGTCTACCACAGCAAGTACAAGCCAAAGAGGCTGGGCAATGACATCGCCCTTATGAAGCT
GGCCGGGCCACTCACGTTCAATGAAATGATCCAGCCTGTGTGCTCGCTCCCACTCTGAAGAGA
ACTTCCCCGATGGAAAAGTGTGCTGGACGTCAGGATGGGGGGCCACAGAGGATGGAGGTGAC
GCCCTCCCCTGTCTGAACCACGCGGCCGTCCCTTTGATTTCCAACAGATCTGCAACCACAG
GGACGTGTACGGTGGCATCATCTCCCCCTCCATGCTCTGCGCGGGCTACCTGACGGGTGGCG
TGGACAGCTGCCAGGGGGACAGCGGGGGGCCCTGGTGTGTCAAGAGAGGAGGCTGTGGAAG
TTAGTGGGAGCGACAGCTTTGGCATCGGCTGCGCAGAGGTGAACAAGCCTGGGGGTGTACAC
CCGTGTCACTCTCTTCTGGACTGGATCCACGAGCAGATGGAGAGAGACCTAAAAACCTGAA
GAGGAAGGGGACAAAGTAGCCACCTGAGTTCTGAGGTGATGAAGACAGCCGATCCTCCCCCT
GGACTCCCCTGTAGGAACCTGCACACGAGCAGACACCCTTGGAGCTCTGAGTTCCGGCACCA
GTAGCAGGCCCGAAAGAGGACCCCTTCCATCTGATTCCAGCACAACTTCAAGCTGCTTTTT
GTTTTTTGTTTTTTTGGAGTGGAGTCTCGCTCTGTTGCCAGGCTGGAGTGCAGTGGCGAAA
TCCTGTCTCACTGCAGCCTCCGCTTCCCTGGTTCAAGCGATTCTCTGCTCAGCTTCCCCA
GTAGCTGGGACCACAGGTGCCCGCCACCACACCCAACTAATTTTTGTATTTTTAGTAGAGAC
AGGGTTTACCATGTTGGCCAGGCTGCTCTCAAACCCCTGACCTCAAATGATGTGCTGCTT
CAGCCTCCACAGTGCTGGGATTACAGGCATGGGCCACCACGCCTAGCCTCAGCTCCTTTT
TGATCTTCACTAAGAACAAAAGAAGCAGCAACTTGCAAGGGCGGCCTTCCCACTGGTCCAT
CTGGTTTTCTCTCCAGGGTCTTGCAAAATTCCTGACGAGATAAGCAGTTATGTGACCTCAG
TGCAAAAGCCACCAACAGCCACTCAGAAAAGACGCACCAGCCAGAAAGTGCAAACTGCAGTC
ACTGCACGTTTTCTCTAGGGACCAGAACCAACCCACCTTTCTACTTCCAAGACTTAT
TTTCACATGTGGGGAGGTTAATCTAGGAATGACTCGTTTAAGGCCTATTTTCATGATTTCTT
GTAGCATTTGGTGCTTGACGTATTATTGTCTTTGATTCAAATAATATGTTTCTTCCCT
CATTGTCTGGCGTGTCTGCGTGGACTGGTGACGTGAATCAAATCATCCACTGAAA

FIGURE 28

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45234

><subunit 1 of 1, 453 aa, 1 stop

><MW: 49334, pI: 6.32, NX(S/T): 1

MGENDPPAVEAPFSFRSLFGLDDLKISPVAPDADAVAAQILSLLPLKFFPIIVIGIILILA
LAIGLGIHFDCSGKYRCRSSFKCIELIARCDGVSDCKDGEDEYRCVVRVGGQNAVLQVFTAAS
WKTMCSDDWKGYANVACAQLGFPSYVSSDNLRVSSLEGQFREEFVSIDHLLPDDKV TALHH
SVYVREGCASGHVVTLQCTACGHRRGYSSRIVGGNMSLLSQWPWQASLQFQGYHLCGGSVIT
PLWIITAAHCVDLYLPKSWTIQVGLVSLLDNPAPSHLVEKIVYHSKYKPKRLGNDIALMKL
AGPLTFNEMIQPVCLPNSEENFPDGKVCWTSGWGATEDGGDASPVLNHAAPLISNKICNHR
DVYGGIISPSMLCAGYLTGGVDSCQGDSSGGLVCQERRLWKLVGATSFSGIGCAEVNKP G VYT
RVTSFLDWIHEQMERDLKT

Signal Peptide:

amino acids 1-20

Transmembrane domain:

amino acids 240-284

FIGURE 29

CCCACGCGTCCGTCTAGTCCC CGGGCCAACTCGGACAGTTTGCTCATTATTGCAACGGTCAAGGCTGGCTTGT
GCCAGAAACGGCGCGCGCGCGCACGACGACACACACGCGGGGAAACTTTTTTAAAAATGAAAGGCTAGAAGA
GCTCAGCGGGCGCGCGGGCGCTGCGCGAGGGCTCCGGAGCTGACTCGCCGAGGCAGGAAATCCCTCCGCTCGCGA
CGCCCGGCCCGGCTCGGCGCCCGCGTGGGATGGTGACGGCTCGCCGCGGGGCCGAGAGCTGCTGCACTGAAG
CGCCGCGACCAATGGCAGCGCGCGCGCTGCCCGTGTCCCGCCCGCGCCCTCTGCTCGCCCTGGCCGGTGCTCT
GCTCGCGCGCTCGGAGCCCGAGGGGTGAGCTTATGGAACCAAGGAAGAGCTGATGAAGTTGTCAGTGGCTCTGT
TCGGAGTGGGACCTCTGGATCCAGTGAAGAGCTTCGACTCCAAGAAATCATCCAGAAGTGTGTAATATTCGACT
ACAACGGGAAAGCAAAGAACTGATCATAAATCTGGAAGAAATGAAGGTCTCATTGCCAGCAGTTTACGGAAC
CCACTATCTGCAAGACGGTACTGATGTCTCCCTCGCTCGAAATTACACGGGTCACTGTTACTACCATGGACATGT
ACGGGGATATTCTGATTACGACAGTCACTCTCAGCACGTGTCTGGTCTCAGGGGACTTATTGTGTTTGAATAATGA
AAGCTATGTCTTAGAACCAATGAAAGTGAACCAACAGATACAAACTCTTCCAGCGAAGAAGCTGAAAAGCGT
CGGGGATCATGTGGATCACATCAACACACCAAACTCGCTGCAAGAATGTGTTCCACCACCTCTCAGAC
ATGGGCAAGAAGGCATAAAAGAGAGACCCTCAAGGCAACTAAGTATGTGGAGCTGGTGATCGTGGCAGACAACCG
AGAGTTTCAGAGGCAAGGAAAAGATCTGGAAGAAAGTTAAGCAGCGATTAAATAGAGATTGCTAATCACGTTGACAA
GTTTTACAGACCACTGAACATTTCGGATCGTGTGTGGTAGGCGTGGAAAGTGTGGAATGACATGGCAAAATGCTCTGT
AAGTCAGGACCCATTCAACAGCCTCCATGAAATTTCTGGACTGGAGGAAGATGAAGCTTCTACTCGCAAAATCCCA
TGACAAATGCGCAGCTGTGTCAGTGGGTTTATTTCGAAGGACCACTCGGACGTGCCCCCAATCATGAGCATGTG
CAGCGCAGCAGCTTGGGGGAATTTGATGGACCACTCAGACCAATCCCTTGGTGGCAGCCGTGACCCCTGGCACA
TGAGCTGGGCCCAATTTTCGGGATGAATCATGACACACTGGACAGGGGCTGTAGCTGTCAAATGGCGGTTGAGAA
AGGAGGCTGCATCATGAACGCTTCCACCGGGTACCATTCCCATGGTGTTCAGCAGTTGACAGCAGGAAGGACTT
GGAGACCGAGCTGGAGAAAGGAATGGGGGTGTGCCGTGTTAACCTGCGGGAAGTCAGGGAGTCTTTCCGGGGCCA
GAAGTGTGGGAACAGATTTGTGGAAGAAGGAGAGAGTGTGACTGTGGGAGGCCAGAGGAATGTATGAATCGCTG
CTGCAATGCCACCTGTACCTTGAAGCCGAGCGCTGTGTGCGCACATGGGCTGTGCTGTGAAGACTGCCAGCT
GAAGCTCGAGGAACAGCGTGCAGGGGACTCCAGCAACTCCTGTGACTCCAGAGTCTCTGCAAGGCGGACGCC
TCACCTGCCAGCCAAATGTGTACTCTGCAGCATGGGCACTCATGTCCAGATGTGGACGGCTACTGCTCAATGGCCAT
CTGCCAGACTCAGCAGCAGCAGTGTGTACGCTCTGGGACCAAGTGTCAAACTGCCCTGGGATCTGCTTTGA
GAGAGTCAATTTCTGCAGGTGATCTTATGGCAACTGTGGCAAGTCTCGAAGAGTTCCTTTGCCAAATGCGAGAT
GAGAGATGCTAAATGTGGAAGAAATCCAGTGTCAAGGAGGTGCCAGCCGGCCAGTCATTGGTACCAATGCCGTTTC
CATAGAAACAAACATCCCTCTGCAGCAAGGAGGCCGGAATTTGTGCCGGGGGACCCACGTGTACTTGGGCGATGA
CATGCGGAGCCAGGGCTTGTGCTTGCAGGCACAAAGTGTGCAGATGGAAAAATCTGCCGTAATCGTCAATGTCA
AAATATTAGTGTCTTTGGGGTTTCAAGAGTGTGCAATGCAGTGCCACGGCAGAGGGGTGTGCAACACAGGAAGAA
CTGCCACTGCGAGGCCCACTGGGCACCTCCCTTCTGTGACAAGTTTGGCTTTGAGGAAGACACAGACAGCGGCC
CATCCGGCAAGCAGAAGCAAGGCAGGAAGCTGCAGAGTCCACAGGGAGCTGGGCCAGGGCCAGGACCCGTGGG
ATCGCAGGAGCATGCGTCTACTGCCTCACTGACACTCATCTGAAGCCCTCCCATGACATGGAGACCGTGACCACTG
CTGCTGCAGAGGAGGTCAAGCGTCCCAAGGCCCTCCTGTGACTGGCAGCATTGACTCTGTGGCTTTGCCATCGTT
TCCATGACAAACAGACCAACACAGTTCTCTGGGGCTCAGGAGGGGAAGTCCAGCTTACCAGGCACGTCTGCAGAAA
CAGTGCAAGGAAGGCCAGCGACTTCTGTGTTGAGCTTCTGTGTAACCATGGACATGCTTCAATGCTGCTCTCTGAG
AGAGTAGCAGGTTTACCACCTCTGGCAGGCCCCAGCCCTGCAGCAAGGAGGAAGAGGACTCAAAGTCTGGCCCTTTC
ACTGAGCTTCCACAGCAGTGGGGGAGAGCAAGGGTTGGGCCAGTGTCCCTTTCCCCAGTGACACCTCAGCCT
TGGCAGCCCTGATGACTGGTCTCTGGCTGCAACTTAATGCTCTGATATGGCTTTTAGCATTATTATATGAAAAAT
AGCAGGGTTTATGTTTATTAATTTATCAGAGACCTGCCACCAATTCATCTCCATCCAAGCAAATCGAATGGCAA
TGAAACAAACTGGAGAAGAGGTAGGAGAAAGGGCGGTGAACTCTGGCTCTTTGCTGTGACATGCGCTGACCCAGC
AGTATCAGGTTTGGGGTTTGCAGAAAGCCAGGGAAACCCACAGAGTCAACAAACCTTCAATTTAAACAGTAAGAA
TGTTAAAAAGTGAACCAATGTAAGAGCCTTAATCCATCCCCGTGGCCATTACTGCATAAAATAGAGTGCAATTT
GAAAT

0097803-6160
10707-608260

FIGURE 30

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA49624

><subunit 1 of 1, 735 aa, 1 stop

><MW: 80177, pI: 7.08, NX(S/T): 5

MAARPLPVSPARALLLALAGALLAPCEARGVSLWNQGRADEVVSA SVRSGDLWIPVKSFD SK
NHPEVLNIRLQRESKELI INLERNEGLIASSFTETHYLQDGT DVSLARNYTGH CYHGHV RG
YSDSAVSLSTCSGLRGLIVFENESYVLEPMKSATNRYKLFPAKKLKSVRGSCGSHHNTPNLA
AKNVFPPPSQTWARRHKRETLKATKYVELVIVADNREFQRQGDLEKVKQRLIETIANHVDKF
YRPLNIRIVLVGVEVWNDMDKCSVSQDPFTSLHEFLDWRKMKLLPRKSHDNAQLVSGVYFQG
TTIGMAPIMSMCTADQSGGIVMDHSDNPLGAAVTLAHELGHNF GMNHD TLDRGCSQMAVEK
GGCIMNASTGYPPFMVFS SCSRKDLET SLEKGMGVCLFNLPEVRESFGGQKCGNRFVEEGEE
CDCGEPEECMNRCNATTCTLKPD AVCAHGLCCEDCQLKPAGTACRDSSNSCDLPEFCTGAS
PHCPANVYLHDGHSCQDVDGYCYNGICQTHEQQCVTLWGPGAKPAPGICFERVNSAGDPYGN
CGKVSKSSFAKCEMRDAKCGKIQCQGGASRPVIGTNAVSIETNIP LQQGGRILCRGTHVYLG
DDMPDPGLVLAGTKCADGKICLN RQCQNISVFGVHECAMQCHGRGVCNNRKNCHCEAHWAPP
FCDKFGFGGSTDSGPIRQAEARQEAAESNRERGGQGEFVGSQEHASTASLTLI

Signal peptide:

amino acids 1-28

125121-1046240

TCCCAAGGCTTCTTGGATGGCAGATGATTNTGGGGTTTTGCATTGTTTTCCCTGACAACGAAA
ACAAAAACAGTTTTGGGGGTTT CAGGAGGGGAANTCCAGCCTACCCAGGAAGTTTGCAGAAACA
GTGCAAGGAAGGGCAGGANNTTCTGGTTGAGNTTTTTGNATAAACATGGACATGNNTTCAGTG
CTGCTCNTGAGAGAGTAGCAGGTTACCACTTTTGGCAGGCCCCAGCCCTGCAGCAAGGAGGA
AGAGGACTCAAAAGTTTGGCCTTTCACTGAGCCTCCACAGCAGTG GGGGGAGAAGCAAGGGTT
GGGCCCAAGTGTCCCTTTTCCCCAGTGACACCTCAGCCTTGGCAGCCCTGATAACTGTTNTNT
GGCTGCAANTTAATGCTNTGATATGGCTTTTAGCATTTATTATATGAAAAATAGCAGGGTTTT
AGTTTTTAATTTATCAGAGACCCTGCCACCCATTCCATNTCCATCCAAG

FIGURE 32

CATCCTGCAACATGGTGAAACCACGCCTGGCTAATTTTGTGTATTTTGGTAGAGATGGGA
TTTCACCGTGTAGCCAGGATTGTCTCAATCTGACCTCATGATCTGCCCCTCGGCCTCCC
AAAGTGCTGGGATTACAGGCGAGTGCAACCACCCCGCCACAACTTTTTAAGAAGTTAAT
GAAACCATACTTTTACATTTTAAATGACAGGAAAATGCTCACAATAATTGTTAACCCAAAA
TTCTGGATACAAAAGTACAATCTTTACTGTGTAAATACATGTATATGTACTATATGAAAATA
TACCAAATATCAATAATACTTTATCTCTGGGTAAAAACCTCTTCTCATACCCCTGTGCTAACAA
CTTTTAACAAAAAATTTGCATCACTTTTAAGAATCAAGAAAAATTTCTGAAGGTCATATGGG
ACAGAAAAAAAACCAAGGAAAAATCACGCCACTTGGGAAAAAAGATTGCAAATCTGCCT
TTTTATAGATTTGTAATTAATAAGGTCCAGGCTTTCTAAGCAACTTAAATGTTTTGTTTCGA
AACAAAGTACTTGTCTGGATGTAGGAGGAAAGGGAGTGATGTCAGTCCATTATGATGCCCC
TTGAATATAAGACCTTACTTGTCTATCTCCCTGCACCAGCCAGGAGCCACCCATCTCCAGC
ACACTGAGCAGCAAGCTGGACACACGGCACACTGATCCAAATGGGTAAGGGGATGGTGGCGA
TGCTCATTCTGGGTCTGCTACTTCTGGCGCTGCTCCTACCCGTGCAGGTTTCTTCATTTGTT
CCTTTAACAGTATGCCGGAAGCTACTGCAGCCGAAACCACAAAGCCCTCCAACAGTGCCCT
ACAGCCTACAGCCGGTCTCCTTGTGGTCTTGCTTGCCTTCTACATCTCTACCATTAAAGAGG
CAGGTCAAGAAACAGCTACAGTTCTCCAACCCATACACTAAAACCGAATCCAAATGGTGCCT
AGAAGTTCAATGTGGCAAGGAAAAAACAGGTCTTCATCAAATCTACTAATTTCACTCCTT
ATTAACAGAGAAACGCTTGAGAGTCTCAAAGTGGACTGGTTTAAAGAGCATCTGAAGGATTT
GACTAGATGATAAATGCCTGTACTCCAGTACTTTGGGAGGCCTAGGCCGGCGGATCACCTG
AGGTCAAGGAGTTTGAGACTAACCTGGCCAAAATGGTGAAACCCCATCTGTACTAAAAATACA
AATATTGACTGGGCGTGGTGGTGAAGTGCCTGTGATCCAGCTACTCAGGTGGCTGAAGCAGG
ACAATCACTTGAAGTCAAGGAGGCAGAGGTTGCAGTGAGCTGAGATCGCGCTACTGCACTCTA
GCCTAGCCTGGGCAACAGAGTGAGACTTCGTCTCAAAAAAAAAAAGCCAAGTGCAGTGGCT
CACGCCTGTAATCCCGGCCTTTGGGAGGCCGAGGTGGGCGGATCACGAGGTCAGGAGATCA
AGACCATCCTGGCTAATACAGTGAAACCCCTGTCTCTACTAAAAATACAAAAAATTAGCCGGG
GATGGTGGCAGGCACCTGGAGTCCAGCTACTCGGAGGCTGAGGCAGGAGAATAGCGTGAA
CTCAGGAGGCGGAGCTTGCACTGAGCCGAGATTGCGCTACTGCACTCCAGCTGGGCGACAG
CGCGAGACTCCGTCTCAAAAAAAAAAAAAAAAAAAAAA

FIGURE 33

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48309

><subunit 1 of 1, 67 aa, 1 stop

><MW: 6981, pI: 7.47, NX(S/T): 0

MGKGMVAMLILGLLLLALLLPVQVSSFVPLTSMPEATAAETTKPSNSALQPTAGLLVVLLAL

LHLYH

Signal peptide:

amino acids 15-27

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 35

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA46776

><subunit 1 of 1, 432 aa, 1 stop

><MW: 47629, pI: 5.90, NX(S/T): 0

MPARPGRLLPLLARPAALTALLLLLLGHGGGGRWGARAQEAAAAAADGPPAADGEDGQDPHS
KHLTYTADMFTTHGIQSAAHFVMFFAPWCGHCQRLQPTWNDLGDKYNSMEDAKVYVAKVDCTAH
SDVCSAQGVRGYPTLKLFKPGQEAVKYQGPRDFQTLNWMQLTLNBEPVTPEPEVEPPSAPE
LKQGLYELSSANFELHVAQGDHFIKFFAPWCGHCKALAPTWEQLALGLEHSETVKIGKVDCT
QHYELCSGNQVRGYPTLLWFRDGGKKVDQYKGKRDLESLERYVESQLQRTETGATETVTPSEA
PVLAAEPEADKGTVLALTENNFDDTIAEGITFIKFYAPWCGHCKTLAPTWEELSKKEFPGLA
GVKIAEVDCTAERNICSKYSVRGYPTLLLFRGGKKVSEHSGGRDLDSLHRFVLSQAKDEL

Signal sequence:

amino acids 1-32

009756403 10:50:01 E0452600

FIGURE 36

CTTTTCTGAGGAACCAAGCAATGAATGGCTTTGCATCCTTGCTTCGAAGAAACCAATTTAT
CCTCCTGGTACTATTTTCTTTTGCAAATTCAGAGTCTGGGTCTGGATATTGATAGCCGTCCTA
CCGCTGAAGTCTGTGCCACACACAATTTACCAGGACCCAAAGGAGATGATGGTGAAGAAA
GGAGATCCAGGAGAAGAGGGGAAAGCATGGCAAAGTGGGACGCATGGGGCCGAAAGGAATTAA
AGGAGAACTGGGTGATATGGGAGATCAGGGCAATATTGGCAAGACTGGGCCCATTTGGGAAGA
AGGGTGACAAAGGGGAAAAAGGTTTGCTTGGAATACCTGGAGAAAAAGGCAAAGCAGGTACT
GTCTGTGATTGTGGAAGATACCGGAAATTTGTTGGACAACCTGGATATTAGTATTGCTCGGCT
CAAGACATCTATGAAGTTTGTCAAGAATGTGATAGCAGGGATTAGGGAACTGAAGAGAAAT
TCTACTACATCGTGCAGGAAGAGAAGAACTACAGGGAATCCCTAACCCACTGCAGGATTCGG
GGTGGAATGCTAGCCATGCCCAAGGATGAAGCTGCCAACACACTCATCGTGACTATGTTGC
CAAGAGTGGCTTCTTTTCGGGTGTTTATTGGCGTGAATGACCTTGAAAGGGAGGGACAGTACA
TGTCCACAGACAACACTCCACTGCAGAACTATAGCAACTGGAATGAGGGGGAACCCAGCGAC
CCCTATGGTCATGAGGACTGTGTGGAGATGCTGAGCTCTGGCAGATGGAATGACACAGAGTG
CCATCTTACCATGTACTTTGTCTGTGAGTTCATCAAGAAGAAAAAGTAACTTCCCTCATCCT
ACGTATTTGCTATTTTCCTGTGACCGTCATTACAGTTATTGTTATCCATCCTTTTTTTTCCTG
ATTGTACTACATTTGATCTGAGTCAACATAGCTAGAAAATGCTAAACTGAGGTATGGAGCCT
CCATCATCAAAAAAAAAAAAAAAAAA

FIGURE 37

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50980

><subunit 1 of 1, 277 aa, 1 stop

><MW: 30645, pI: 7.47, NX(S/T): 2

MNGFASLLRRNQFILLVLFLQLQISLGLDIDSRPATAEVCATHTISPGRP KDDGEKGDPEEG
KHGKVGRMGPKGIKGE LGDMGDQGNIGKTGP IGGKGDKEKGLLGIPGEK GKAGTVDCGRY
RK FVGQLDISIARLKTSMKFVK NVIAGIRETEEFYIIVQEEKNYRESLTHCRIRGGMLAMP
KDEAANTLIADYVAKSGFFRVF IGVNDLEREGQYMSTDNTP LQNYSNWNEGEPSDPYGHEDC
VEMLSSGRWNDTECHLTMYFVCEFIKKKK

Signal peptide:

amino acids 1-25

109701-109702-109703

FIGURE 38

GGTTCTATCGATTCGAATTCGGCCACACTGGCCGGATCCTCTAGAGATCCCTCGACCTCGAC
CCACGCGTCCGCTGCTCTCCGCCCGTGTGGAGTGGTGGGGGCTGGGTGGGAAATGGGCGTGT
GCCAGCGCACGCGCGCTCCCTGGAAGGAGAAGTCTCAGCTAGAACGAGCGGCCCTAGGTTTT
CGGAAGGGAGGATCAGGGATGTTTGCAGCGCGCTGGAACAGACGGTGCCGATAGAGGAAGC
GGGCTCCATGGCTGCCCTCCGCTGCTGCTGCCCTGCTGCTGTTGCTACCGCTGCTGCTGCTGA
AGCTACACCTCTGGCCGCGAGTTGCGCTGGCTTCCGGCGGAACTTGGCCTTTGCGGTGCGAGCT
CTGTGCTGCAAAAAGGCTCTTCGAGCTCGCGCCCTGGCCGCGGCTGCCGCCGACCCGGAAGG
TCCCAGGGGGGCTGCAGCTGGCCTGGCGCTCGCGAACTGGCCAGCAGCGCGCCGCGC
ACACCTTTCTCATTACGGCTCGCGCGCTTTAGCTACTCAGAGGCGGAGCGCGAGAGTAAC
AGGGCTGCACGCGCTTCTCTACGTGCGCTAGGCTGGGACTGGGACCCGACGCGCGCGACAG
CGGCGAGGGGAGCGCTGGAGAAGGCGAGCGGGCAGCGCCGGGAGCCGGAGATGCAGCGGCCG
GAAGCGCGCGGAGTTTGCCGGAGGGGACGTTGCCGCCAGAGGTGGAGGAGCCGCCGCCCT
CTGTACCTGGAGCAACTGTGGCGTCTCTCTCCCGCTGGCCAGAGTTTCTGTGGCTCTG
GTTTCGGGCTGGCCAAAGGCGCGCTGCGCACTGCTTTGTGCCACCGCCTTGCGCCGGGGCC
CCCTGCTGCATGCTTCCGCGAGCTGCGGCGCGCGCTGTGTGCTGGCGCCAGAGTTTCTG
GAGTCCCTGGAGCCGGAACCTGCCCGCCCTGAGAGCCATGGGGCTCCACCTGTGGGCTGCAGG
CCCAGGAACCCACCTGCTGGAATTAGCGATTTGCTGGCTGAAGTGTCCGCTGAAGTGGATG
GGCCAGTGCCAGGATACCTCTTCCCCCAGAGCATAACAGACAGTGCCCTGTACATCTTC
ACCTCTGGCACCAAGGCTTCCCAAGGCTGCTCGGATCAGTCACTGAAGATCCTGCAATG
CCAGGGCTTCTATCAGCTGTGTGGTGTCCACCAGGAAGATGTGATCTACCTCGCCCTCCAC
TCTACCACATGTCCGGTTCCTGCTGGGCATCGTGGGCTGCATGGGCATTGGGGCCACAGTG
GTGCTGAAATCCAAGTTCTCGCTGCTCAGTTCTGGGAAGATTGCCAGCAGCACAGGGTGAC
GGTGTTCAGTACATTGGGGAGCTGTGCCGATACCTTGTCAACCAGCCCCGAGCAAGGCAG
AACGTGGCCATAAGGTCCGGCTGGCAGTGGGCAGCGGGCTGCGCCAGATACCTGGGAGCGT
TTTGTGCGGCGCTTCCGGCCCTGCAAGTGTGGAGACATATGGACTGACAGAGGGCAACGT
GGCCACCATCAACTACACAGGACAGCGGGCGCTGTGGGGCGTGCTTCTGGCTTTACAAGC
ATATCTTCCCTTCTCCTTGATTGCTATGATGTCAACACAGGAGAGCCAATTCGGGACCCC
CAGGGCACTGTATGGCCACATCTCCAGGTGAGCCAGGGCTGCTGGTGGCCCCGGTAAGCCA
GCAGTCCCCATTCTGGGCTATGCTGGCGGGCAGAGCTGGCCAGGGGAAGTTGCTAAAGG
ATGTCTTCCGGCCTGGGATGTTTTCTTCAACACTGGGGACCTGCTGGTCTGCGATGACCAA
GGTTTTCTCCGCTTCCATGATCGTACTGGAGACACCTTCAGGTGGAAGGGGAGAAATGTGGC
CACAAACCGAGGTGGCAGAGGTCTTCGAGGCCCTAGATTTCTTCAGGAGGTGAACGTCTATG
GAGTCACTGTGCCAGGGCATGAAGGCAGGGCTGGAATGGCAGCCCTAGTTCTGCGTCCCCC
CAGCTTTGGACCTTATGCAGCTCTACACCCAGCTGTCTGAGAACTGCCACCTTATGCCCC
GCCCGATTCTCAGGCTCCAGGAGTCTTTGGCCACCACAGAGACCTTCAAACAGCAGAAAG
TTCGATGGCAAATGAGGGCTTCGACCCAGCACCTGTCTGACCCACTGTACGTTCTGGAC
CAGGCTGTAGGTGCCCTACCTGCCCTCACAACTGCCCGGTACAGCGCCCTCCTGGCAGAAA
CCTTCGAATCTGAGAACTTCCACCTGAGGCACCTGAGAGAGGAACCTGTGGGGTGGGG
CCGTGTCAGTGTACTGGGCTGTACGGGATCTTTCTATACCAGAACTGCGGTCACTATTTT
GTAATAAATGTGGCTGGAGCTGATCCAGCTGTCTGACCTAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAGGGCGGCCGACTCTAGAGTCGACCTGCAGTAGGGATAACAGGGTAATAAGC
TTGGCGCCATGGCCCAACTTGTTTATTGCAG

FIGURE 39

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50913
><subunit 1 of 1, 730 aa, 1 stop
><MW: 78644, pI: 7.65, NX(S/T): 2
MGVCQRTRAPWKEKSQLERAAALGFRKGGSGMFASGWNQTVPIEEAGSMAALLLLPLLLLLPL
LLLKLHLWPQLRWLPADLAFAVRALCCKRALRARALAAAAADPEGPEGGCSLAWRLAELAQQ
RAAHTFLIHGSRFFSYSEAESERENRAARAFLRALGWDWGPDDGDSGEGSAGEGERAAPGAGD
AAAGSGAEFAGGDGAARGGGAAPLSPGATVALLLPAGPEFLWLWFLAKAGLRTAFVPTAL
RRGPLLHCLRSCGARALVLAPEFLESLEPDLPALRAMGLHLWAAGPGTHPAGISDLLAEVSA
EVDGPVPGYLSSPQSITDTCLYIFTSGTTGLPKAARISHLKILQCQGFYQLCGVHQEDVIYL
ALPLYHMSGSLLGIVGCMGIGATVVLKSKFSAGQFWEDCQQHRVTVFQYIGELCRYLVNQPP
SKAERGHKVRLLAVGSGLRPDTWERFVRRFGPLQVLETYGLTEGNVATINYTGQRGAVGRASW
LYKHIFPFSLIRYDVTTGEPIRDPOGHCMATSPGEPGLLVAPVSQQSPFLGYAGGPFLAQGK
LLKDVFRPGDVFFNTGDLVLCDDQGFLRFHDRTGDTFRWKGENVATTEVAEVFEALDFLQEV
NVYGVTVPGHEGRAGMAALVLRPPHALDLMQLYTHVSENLPYARPRFLRLQESLATTETFK
QQKVRMANEGFDPSTLSDEPLYVLDQAVGAYLPLTTARYSALLAGNLR
```

Type II transmembrane domain:

amino acids 45-65

Other transmembrane domain:

amino acids 379-398

cAMP- and cGMP-dependent protein kinase phosphorylation site
starting at amino acid 136

CUB domain protein motif

amino acids 254-261

putative AMP-binding domain signature

amino acids 332-343

N-glycosylation sites

amino acids 37-40 and 483-486

FIGURE 40

CCTGTGTTAAGCTGAGGTTTCCCTAGATCTCGTATATCCCCAACACATACCTCCACGCACA
CACATCCCCAAGAACCTCGAGCTCACACCAACAGACACACGCGCGCATACACACTCGCTCTC
GCTTGTCCATCTCCCTCCCGGGGAGCGCGCGCTCCACCTTTGCCGCACACTCCGGC
GAGCCGAGCCCGCAGCGCTCCAGGATTCTGCGGCTCGGAAGCTCGGATTGCAGCTCTGAACCC
CCATGGTGGTTTTTAAACACTTCTTTTCTTCTCTCTCGTTTGGATTGACCCGTTTCCA
TCTGGGGGCTAGAGGAGCAAGGCAGCAGCCTTCCAGCCAGCCCTTGTGGCTTGCCATCGT
CCATCTGGCTTATAAAAGTTTGTCTGAGCGCAGTCCAGAGGGCTCGGCTGCTCGTCCCTCGG
CTGGCAGAAGGGGTGACCTGGCGAGCGCGAGGAGCGCGCTGCCTTGGCGGGCTTT
CGGCTTGAGGGGCAAGGTGAAGAGCGACCCGGCCGTGGGGTTTACCAGAGCTGGATTTGTATG
TTGCACCAATGCCCTTCTTGGATCGGGGCTGTGATTCTTCCCTCTTGGGGCTGCTGCTCTCC
TCCCCGCCGGGGCGGATGTGAAGGCTCGGAGCTGCGGAGAGGTCCGCCAGGCGTACGGTGCC
AAGGGATTACAGCTGGCGGACATCCCCACCAGGAGATCGCAGGGGAACACTTAAGAATCTG
TCCTCAGGAATATACATGCTGCACCACAGAATGGAAGACAGTTAAGCCAACAAAGCAAAC
TCGAATTTGAAAACCTTGTGGAAGAGACAAGCCATTTTGTGCGCACCACTTTTGTGTCCAGG
CATTAAGAAATTTGACGAATTTTTCGAGAGCTCCTGGAGAATGCAGAAAAGTCACTAAATGA
TATGTTTTGTACGGACCTATGGCATGCTGTACATGCAGAATTCAGAAGTCTTCCAGGACCTCT
TCACAGAGCTGAAAAGGTACTACACTGGGGGTAAATGTGAATCTGGAGGAAATGCTCAATGAC
TTTTGGGCTCGGCTCCTGGAACGGATGTTTCAGCTGATAAACCTCAGTATCACTTCAGTGA
AGACTACCTGGAATGTGTGAGCAAATACTAGACCAGCTCAAGCCATTTGGAGACGTGCCCC
GGAAACTGAAGATTACAGTTACCCGCGCCTTCATTGCTGCCAGGACCTTTGTCCAGGGGCTG
ACTGTGGGCAGAGAAGTTGCAAACCGAGTTTCCAAGGTACGCCAACCCAGGGTGTATCCG
TGCCCTCATGAAGATGCTGTACTGCCATACTGTGCGGGGCTTCCACTGTGAGGCCCTGCA
ACAACCTACTGTCTCAACGTCATGAAGGGCTGCTTGGCAAAATCAGGCTGACCTCGACACAGAG
TGGAACTCTGTTTATAGATGCAATGCTCTTGGTGGCAGAGCGACTGGAGGGGCCATTCAACAT
TGAGTCGGTCATGGACCCGATAGATGTCAAGATTTCTGAAGCCATTATGAACATGCAAGAAA
ACAGCATGCAGGTGTCTGCAAAGGTCTTTCAGGGATGTGGTCAGCCCAAACCTGCTCCAGCC
CTCAGATCTGCCCGCTCAGCTCCTGAAAATTTTAAATACACGTTTCAGGCCCTACAATCCTGA
GGAAAGACCAACAACCTGCTGCAGGCACAAGCTTGGACCGGCTGGTCACAGACATAAAAGAGA
AATTGAAGCTCTCTAAAAAGGTCTGGTCAGCATTACCCTACACTATCTGCAAGGACGAGAGC
GTGACAGCGGGCAGCTCCAACGAGGAGGAATGCTGGAACGGGCACAGCAAAGCCAGATACTT
GCCTGAGATCATGAATGATGGGCTCACCAACCAGATCAACAATCCCGAGGTGGATGTGGACA
TCACTCGGCCGTGACACTTTTCATCAGACAGCAGATTATGGCTCTCCGTGTGATGACCAACAAA
CTAAAAAACGCCTACAATGGCAATGATGTCAATTTCCAGGACACAAGTGATGAATCCAGTGG
CTCAGGGAGTGGCAGTGGGTGCATGGATGACGCTGTGTCCCACGGAGTTTGAGTTTGTACCA
CAGAGGCCCCCGCAGTGGATCCCGACCGGAGAGAGGTGGACTCTTCTGCAGCCAGCGTGGC
CACTCCCTGCTCTCCTGGTCTCTCACCTGCATTGTCTTGGCACTGCAGAGACTGTGCAGATA
ATCTTGGGTTTTTGGTCAGATGAACTGCATTTTAGCTATCTGAATGGCCAACTCACTTCTT
TTCTTACACTCTTGGCAATGGACCATGCCACAAAAATTTACCCTTTCTATGAGAAGAGAG
CAGTAATGCAATCTGCCTCCCTTTTGGTTTTCCCAAAGAGTACGGGTGGCCAGACTGAACTG
CTTCCTCTTCTCTCAGCTATCTGTGGGACCTTGTTTATTCTAGAGAGAAATCTTACTCAA
ATTTTTCGTACCAGGAGATTTTCTTACCTTCATTTGCTTTTATGTGTCAGAAGTAAAGGAAT
CTCACGTTGTGAGGGTTTTTTTTTCTCATTTAAAT

FIGURE 41

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50914

><subunit 1 of 1, 555 aa, 1 stop

><MW: 62736, pI: 5.36, NX(S/T): 0

MPSWIGAVILPLLGLLLSLPAGADVKARSCGEVRQAYGAKGFSLADIPYQEIAGEHLRICPQ
EYTCCTTEMEDKLSQQSKLEFENLVEETSHFVRTTFVSRHKKFDEFFRELLENAEKSINDMF
VRTYGMMLYMQNSEVFQDLFTTELKRYYTGGNVNLEEMLNDFWARLLERMFQLINPQYHFSEDY
LECVSKYTDQLKPFQDVPRKLKIQVTRAFIAARTFVQGLTVGREVANRVSKVSPTPGCIRAL
MKMLYCPYCRGLPTVRPCNNYCLNVMKGCLANQADLDTEWNLFIDAMLLVAERLEGPFNIES
VMDPIDVKISEAIMNMQENSMQVSAKVFQCGGQPKPAPALRSARSAPENFNTRFRPYNPEER
PTTAAGTSLDRLVTDIKEKLKLSKKVWSALPYTICKDESVTAGTSNEEECWNGHSKARYLPE
IMNDGLTNQINNPEVDVDITRPDTFIRQQIMALRVMTNKLKNAYNGNDVNFQDTSDESSGSG
SGSGCMTDDVCPTFEFEFVTTEAPAVDPPDRREVDSSAAQRGHSLLSWSLTCIVLALQRLCR

Signal peptide:

amino acids 1-23

FIGURE 42A

CGGACGCGTGGGCGGACGCGTGGGCAAAGAACTCGGAGTGCCAAAGCTAAATAAGTTAGCTGAGAAAAACGCACG
CAGTTTGCAGCGCCTGCGCCGGGTGCGCCAACTACGCAAGACCAAGCGGGCTCCGCGCGGACCGGCCGCGGGGC
TAGGGACCCGGCTTTGGCCTTCAGGCTCCCTAGCAGCGGGGAAAAGGAATTGCTGCCCGAGTTTCTGCGGAGGT
GGAGGGAGATCAGGAAACGGCTTCTTCTCACTTCGCGCGCTGGTGAGTGTGCGGGAGATTGGCAAACGCCTAGG
AAAGGACTGGGAAAATAGCCCTGGGAAAGTGAGAAAGTGATCAGGAGGCGGTCCACTACGGCAGTTTATCTG
TCTGATCAGAGCCAGACGCGACGCGTCCACTTCGCAGTTCTTTCAGGTGTGGGGACCGCAGGACAGACGCGCCGA
TCCCGCCGCCCTCCGTACCAGCACTCCCAGGAGAGTCAGCCTCGCTCCCAACGTCGAGGGCGCTCTGCGCCACGA
AAAGTTCCTGTCCACTGTGATTCTCAATTCCTTGCTTGGTTTTTTCTCCAGAGAACTTTTGGGTGGAGATATTA
ACTTTTCTTTTCTTTTCTTCTTCTTGGTGGAAGCTGCTCTAGGGAGGGGGAGGAGGAGGAGAGAAAGTGAATGTGC
TGGAGAAGAGCGAGCCCTCCTGTCTTCTCGGAGTCCCATCCATTAAAGCCATCACTTCTGGAAGATTAAAGTTGT
CGGACATGTTGACAGCTGAGAGGAGAGGAGGATTCTTGGCCAGGTGGAGAGTCTTCAACCGTCTGTTGGGTGCATG
TGTGCGCCCGCAGCGGCGCGGGGCGCGTGGTTCTCCGCGTGGAGTCTCACCTGGGACCTGAGTGAATGGCTCCCA
GGGCTGTGTCGGGGCATCCGCTCTCGCTTCTCCACAGGCTGTGTCTGTCTTGGAAAGATGCTAGCAATGGGGG
CGCTGGCAGGATTCTGGATCTCTGCTCCTCACTTATGTTACCTGTCTGGGGCCAGGCTTGAAGAGGAGGG
AAGAAGGGGCTTACTAGCTCAAGCTGGAGAGAACTAGAGCCAGCACAACCTTCCACCTCCAGCCCCATCTCA
TTTTCATCTAGCGATGATCAGGGATTAGAGATGTGGGTTACACGGATCTGAGATTAAACACCTACTCTTG
ACAAGCTCGCTGCCAAGGAGTTAACTGGAGAACTACTATGTCCAGCCTATTTGCACACCATCCAGGAGTCAGT
TTATTATGGAAAGTATCAGATACACCCGGACTTCAACATTCTATCATAGAGCACTACCCAAACCAACTGTTTAC
CTCTGGACAATGCCACCTACCTCAGAACTGAAGGAGGTGGATATTCAACGCATATGGTCGGAAAAATGGCACT
TGGGTTTTAACAGAAAAGAATGCATGCCACCAGAGAGGATTGATACCTTTTTTGGTTCCTTTTTGGGAAGTG
GGGATTACTATACACTACAAATGTGACAGTCTGGGATGTGTGGCTATGACTTGTATGAAAACGACAATGCTG
CCTGGGACTATGACAATGGCATATACTCCACAGATGTACACTCAGAGAGTACAGCAAACTCTTAGCTTCCATA
ACCCACAAAGCCTATATTTTTATATACTGCCTATCAAGCTGTTCACTTCCACCTGCAAGCTCCTGGCAGGTATT
TCGAACACTACCGATCCATTATCAACATAAACAGGAGAAGATATGCTGCCATGCTTCTCTGCTTAGATGAAGCAA
TCAACAACGTGACATTGGCTCTAAAGACTTATGGTTCTATATAACAACAGCATTTATCAITTTACTCTTCAGATAATG
GTGGCCAGCCTACGGCAGGAGGGAGTAACTGGCCTCTCAGAGGTAGCAAAGGAACATATTGGGAAGGAGGGATCC
GGGCTGTAGGCTTTGTGCATAGCCCACTTCTGAAAAACAAGGGAACAGTGTGTAAAGAACTTGTGCACATCACTG
ACTGGTACCCCACTCTCATTTCACTGGCTGAAGACAGATTGATGAGGACATTCAACTAGATGGCTATGATATCT
GGGAGACCATAAGTGAGGCTCTCGCTCACCCCGAGTAGATATTTGCATAACATTGACCCCTATACACCAAGGC
AAAAATGGCTCCTGGGCAGCAGGCTATGGGATCTGGAAACATGCAATCCAGTCAGCCATCAGAGTGCAGCACTG
GAAATTGCTTACAGGAATCTGGCTACAGCGACTGGGTCCCCCTCAGTCTTTTTCAGCAACCTGGGACCGAACCG
GTGGCACAATGAACGGATCACCTTGTCACTGGCAAAAGTGATGGCTTTTCAACATCACAGCCGACCCATATGA
GAGGCTGGACCTATCTAACAGGTATCCAGGAATCGTGAAGAAGCTCTCAGGAGGCTCTCAGGTTCAACAAAG
TGCAGTGCCGCTCAGGTATCCCCCAAAGACCCAGAAGTAACCTAGGCTCAATGGAGGGGTCTGGGGACCATG
GTATAAAGAGGAAACCAAGAAAAAGAGCCAAAGCAAAAATCAGGCTGAGAAAAAGCAAAAGAAAAAGCAAAAAA
GAAGAAGAAACAGCAGAAAGCAGTCTCAGGTAAACAGCAAAATTTGGCTGATAATATCGCTGGCCTAAGCGTCA
GGCTTGTTTTATGCTGTGCCACTCCAGAGACTTCTGCCACTGGCCGCCACACTGAAAACCTGTCTGCTCAGTG
CCAAGTGCTACTCTGCAAGCCACACTTAGAGAGAGTGGAGATGTTATTCTCTCGCTCTTTAGAAAAACGTG
GTGAGTCTGTAGTTCCTGCTGTGCTTCACTCACTGACCAAAACACTGCTTTGAATTATAGGAGGAGAACAATA
ACCTACCATCCGCAAGCATGCTAATTTGATGGAAGTTACAGGGTAGCATGATTAACCTACCTTTGATAAATTA

[illegible]

FIGURE 43

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48296
><subunit 1 of 1, 515 aa, 1 stop
><MW: 56885, pI: 6.49, NX(S/T): 5
MAPRCAGHPPPPSPQACVCPGKMLAMGALAGFWILCLLTYGYLSWGQALEEEEEEGALLAQA
GEKLEPSTTSTSQPHLIFILADDQGFQDVGYHGSEIKTPTLQKLAAGVKLENYYVQPICTP
SRSQFITGKYQIHTGLQHSIIIRPTQPNCLPLDNATLPQKLKEVGYSTHMGVKGWHLGFNRKEC
MPTRRGFDTPFGSLLGSGDYIYTHYKCDSPGMCYDLYENDNAAWDYDNGIYSTQMYTQQRVQQ
ILASHNPTKPIFLYTAYQAVHSPLQAPGRYFEHYRSIININRRRYAAMLSCLEAINNVTLA
LKTGYGFYNNSSIIYSSDNGGQPTAGGSNWPLRGSKGTWEGGIRAVGVFHSPLLKNKGTVCK
ELVHITDWDYPTLISLAEGQIDEDIQLDGYDIWETISEGLRSPRVDILHNIDPYTPRQKMAPG
QQAMGSGTLQSSQPSECSTGNCLQEILATATGSPLSLSATWDRGTGGTMNGSPCQLAKVYGFS
TSQPTHMRGWYLTGTQIES
```

Important Features:

Signal Peptide:

amino acids 1-37

Sulfatases signature 1.

amino acids 120-132

Sulfatases signature 2.

amino acids 168-177

Tyrosine kinase phosphorylation site.

amino acids 163-169

N-glycosylation sites.

amino acids 157-160, 306-309 and 318-321

FIGURE 44

CGGACGCGTGGGTGCGAGTGGAGCGGAGACCCGAGCGGCTGAGGAGAGAGGAGGCGCGGC
TTAGCTGCTACGGGGTCCGGCCGGCGCCCTCCCGAGGGGGCTCAGGAGGAGGAAGGAGGAC
CCGTGCGAGAATGCCCTCTGCCCTTGGAGCCTTGCCTCCCGCTGCTGCTCTCCTGGGTGGCAG
GTGGTTTTCGGGAACGCGGCCAGTGCAAGGCATCACGGGTTGTTAGCATCGGCACGTACAGCCT
GGGGTCTGTCACTATGGAACATAAAGTGGCCTGCTGCTACGGCTGGAGAAGAAACAGCAAGGG
AGTCTGTGAAGCTACATGCGAACCTGGATGTAAGTTTGGTGAAGTGGTGGGACCAACAAT
GCAGATGCTTTCCAGGATACACCGGGAAAACCTGCAGTCAAGATGTGAATGAATGTGGAATG
AAACCCCGCCATGCCAACACAGATGTGTGAATACACCGGAAGCTACAAGTGCCTTTTGCCT
CAGTGGCCACATGCTCATGCCAGATGCTACGTGTGTGAAGTCTAGGACATGTGCCATGATAA
ACTGTCAGTACAGCTGTGAAGACACAGAAGAAGGGCCACAGTGCCTGTGTCCATCCTCAGGA
CTCCGCTTGGCCCCAAATGGAAGAGACTGTCTAGATATTGATGAATGTGCCTCTGGTAAAGT
CATCTGTCCCCTACAATCGAAGATGTGTGAACACATTTGGAAGCTACTACTGCAAATGTCACA
TTGGTTTTCGAAGTGAATATATCAGTGGACGATATGACTGTATAGATATAAATGAATGTACT
ATGGATAGCCATACGTGCAGCCACCATGCCAATTGCTTCAATACCCAAGGGTCCCTCAAGTG
TAAATGCAAGCAGGGATATAAAGGCAATGGACTTCGGTGTCTGCTATCCCTGAAAATTTCTG
TGAAGGAAGTCTCAGAGCACCTGGTACCATCAAAGACAGAATCAAGAAGTTGCTTGCTCAC
AAAAACAGCATGAAAAAGAGGCAAAAATTAAAAATGTTACCCAGAACCCACCAGGACTCC
TACCCCTAAGGTGAAGTTCAGCCCTTCAACTATGAAGAGATAGTTTCCAGAGCGGGAAT
CTCATGGAGGTAAAAAAGGGAATGAAGAGAAAATGAAGAGGGGCTTGAGGATGAGAAAAGAG
AAGAGAAAGCCCTGAAGAATGACATAGAGGAGCGAAGCCTGCGAGGAGATGTGTTTTCCCT
AAGGTGAATGAAGCAGGTGAATTCGGCCTGATTCTGGTCCAAAGGAAAGCGCTAACTTCCAA
ACTGGAACATAAAGATTTAAATATCTCGGTGACTGCAGCTTCAATCATGGGATCTGTGACT
GGAAACAGGATAGAGAAGATGATTTTGACTGGAATCCTGCTGATCGAGATAATGCTATTGGC
TTCTATATGGCAGTTCCGGCCTTGGCAGGTCAACAAGAAAGACATTGGCCGATTGAAACTTCT
CCTACCTGACCTGCAACCCCAAGCAACTTCTGTTTGCTCTTTGATTACCGGCTGGCCGGAG
ACAAAGTGGGAAACTTCGAGTGTGTTGTGAAAAACAGTAACAATGCCCTGGCATGGGAGAAG
ACCACGAGTGAGGATGAAAAGTGAAGACAGGGAAAATTGAGTTGTATCAAGGAACTGATGC
TACCAAAAGCATCATTTTTGAAGCAGAACGTGGCAAGGGCAAAACCGGCGAAATCGCAGTGG
ATGGCGTCTTGCTTGTTTCAGGCTTATGTCCAGATAGCCTTTTATCTGTGGATGACTGAATG
TTACTATCTTTATATTTGACTTTGTATGTCAAGTCCCTGGTTTTTTTGATATTGCATCATAG
GACCTCTGGCATTTTAGAAATTAAGCTGAAAAATTGTAATGTACCAACAGAAATATTATTG
TAAGATGCCTTTCTGTATAAGATATGCCAATATTGCTTTAAATATCATATCACTGTATCT
TCTCAGTCATTTCTGAATCTTCCNCATTATATTATAAATNTGGAAANGTCAGTTTATCTC
CCCTCCTCNGTATATCTGATTTGTATANGTANGTTGATGNGCTTCTCTCTACAACATTTCTA
GAAAAATAGAAAAAAGCACAGAGAAATGTTAACTGTTTGACTCTTATGATCACTTCTTGGG
AACTATGACATCAAAGATAGACTTTTGCCTAAGTGGCTTAGCTGGGTCTTTATAGCCAAAC
TTGTATATTTAATTTCTTTGTAATAATAA

109701-1010260

FIGURE 45

MPLPWSLALPLLLSWVAGGFGNAASARHHGLLASARQPGVCHYGTKLACCYGWRNRNSKGVCE
ATCEPGCKFGECVGPKNKCRCFPGYTGKTCSDVNECGMKPRPCQHRVCVNTHGSKYKCFCLSGH
MLMPDATCVNSRTCAMINCQYSCEDTEEGPQCLCPSSGLRLAPNGRDCLDIDECASGKVICP
YNRRVCVNTFGSYYCKCHIGFELQYISGRYDCIDINECTMDSHTCSSHANCFNTQGSFKCKCK
QGYKGNGLRCSAIPENSVKEVLRAPGTIKDRIKKLLAHKNSMKKKAKIKNVTPPTPTPK
VNLQPFNYEEIVSRGGNSHGGKKGNEEK

Signal peptide:

amino acids 1-21

EGF-like domain cysteine pattern signature.

amino acids 80-91

Calcium-binding EGF-like domains

amino acids 103-124, 230-251 and 185-206

FIGURE 46

GGGAGCTGCTGCTGTGGCTGCTGGTGTGTGCGCGCTGCTCCTGCTCTTGGTGCAGCTGCTG
CGCTTCCTGAGGGCTGACGCGACCTGACGCTACTATGGGCCGAGTGGCAGGGACGACGCCC
AGAATGGGAGCTGACTGATATGGTGGTGTGGGTGACTGGAGCCTCGAGTGGAATTGGTGAGG
AGCTGGCTTACCAGTTGTCTAAACTAGGAGTTTCTCTTGTGTGTGTCAGCCAGAAGAGTGCAT
GAGCTGGAAGGGGTGAAAGAAGATGCCCTAGAGAATGGCAATTTAAAGAAAAAGATATACT
TGTTTTGCCCTTGACCTGACCGACACTGGTTCCTCATGAAGCGGCTACCAAAGCTGTTCTCC
AGGAGTTTGGTAGAATCGACATTCTGGTCAACAATGGTGGAATGTCCAGCGTTCTCTGTGC
ATGGATACCAGCTTGGATGTCTACAGAAAGCTAATAGAGCTTAACTACTTAGGGACGGTGTC
CTTGACAAAATGTGTTCTGCCTCACATGATCGAGAGGAAGCAAGGAAAGATTGTTACTGTGA
ATAGCATCCTGGGTATCATATCTGTACCTCTTTCATTGGATACTGTGCTAGCAAGCATGCT
CTCCGGGGTTTTTTTAATGGCCTTCGAACAGAACTTGCCACATACCCAGGTATAATAGTTTC
TAACATTTGCCCAGGACCTGTGCAATCAAATATTGTGGAGAATTCCCTAGCTGGAGAAGTCA
CAAAGACTATAGGCAATAATGGAGACCAGTCCACAAGATGACAACCAGTCGTTGTGTGCGG
CTGATGTTAATCAGCATGGCCAATGATTTGAAAGAAGTTTGGATCTCAGAACAACCTTTCTT
GTTAGTAACATATTTGTGGCAATACATGCCAACCTGGGCCTGGTGGATAACCAAGATGG
GGAAGAAAAGGATTGAGAACTTTAAGAGTGGTGTGGATGCAGACTCTTCTTATTTTAAATC
TTTAAGACAAAACATGACTGAAAAGAGCACCTGTACTTTTCAAGCCACTGGAGGGAGAAATG
GAAAACATGAAAACAGCAATCTTCTTATGCTTCTGAATAATCAAAGACTAATTTGTGATTTT
ACTTTTAAATAGATATGACTTTGCTTCCAACATGGAATGAAATAAAAAATAAATAATAAAG
ATTGCCATGAATCTTGCAAAA

FIGURE 47

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA36343

><subunit 1 of 1, 289 aa, 1 stop

><MW: 32268, pI: 9.21, NX(S/T): 0

MVVVVTGASSGIGEELAYQLSKLGVSLVLSARRVHELERVKRRCLEGNLKEKDILVLPLDL
TDTGSHEAATKAVLQEFGRIDILVNNGGMSQSRSLCMDTSLDVYRKLIELNYLGTVSLTKCVL
PHMIERKQGKIVTVNSILGIISVPLSIGYCASKHALRGFFNGLRTELATYPGIIVSNICPGP
VQSNIVENSLAGEVTKTIGNNGDQSHKMTTSRCVRLMLISMANDLKEVWISEQPFLLVTYLW
QYMPTWAWWITNKMGGKRRIENFKSGVDADSSYFKIFKTKHD

Important Features:

Signal Peptide:

amino acids 1-31

Transmembrane domain:

amino acids 136-157

Tyrosine kinase phosphorylation site.

106-113 and 107-114

Homologous region to Short-chain alcohol dehydrogenase

amino acids 80-90, 131-168, 1-13 and 176-185

FIGURE 48

GCGACGTGGGCACCGCCATCAGCTGTTGCGCGCTCTTCTCCTCCAGGTGGGGCAGGGGTTTC
GGGCTGTTGGAGCATGTGCTGGGACAGGACAGCATCCTCAATCAATCCAACAGCATATTCCG
TTGCATCTTCTACACACTACAGCTATTGTTAGGTGGCTGCGGACACGCTGGGCCTCTGTCC
TGATGCTGCTGAGCTCCCTGGTGTCTCTCGCTGGTTCTGTCTACTGGCCTGGATCCTGTTT
TTCGTGCTCTATGATTTCTGCATTGTTGTATCACCACCTATGCTATCAACGTGAGCCTGAT
GTGGCTCAGTTTTCCGGAAGGTCGAAGAACCCAGGGCAAGGCTAAGAGGCACTGAGCCCTCA
ACCCAAGCCAGGCTGACCTCATCTGCTTTGCTTTGGTCTTCAAGCCGCTCAGCGTGCCTGTG
GACAGCGTGGCCCCCGCCCCCAAGCCTCAGGAGGGCAACACAGTCCCTGGCGAGTGGCCC
TGGCAGGCCAGTGTGAGGAGGCAAGGAGCCACATCTGCGAGCGGCTCCCTGGTGGCAGACAC
CTGGGTCTCACTGCTGCCCACTGCTTTGAAAAGGCAGCAGCAACAGAATGAATTCTGGT
CAGTGGTCTCTGGGTCTCTGTCAGCGTGAGGGACTCAGCCCTGGGGCCGAAGAGTGGGGGTG
GCTGCCCTGCAAGTTGCCAGGGCTATAACCACTACAGCCAGGGCTCAGACCTGGCCCTGCT
GCAGCTCGCCACCCACGACCCACACACCCCTCTGCCTGCCCCAGCCCGCCCATCGCTTCC
CCTTTGGAGCCTCCTGCTGGGCCACTGGCTGGGATCAGGACACCAAGTGTATGCTCTCTGGGACC
CTACGCAATCTGCGCCTGCGTCTCATCAGTCGCCCCACATGTAACGTATCTACAACCAAGCT
GCACCAAGCAGACCTGTCTCAACCCGCGCCCGGCTGGGATGCTATGTGGGGGCCCCAGCCTG
GGGTGCAGGGCCCTGTGAGGGAGATTCCGGGGGCCCTGTGCTGTGCCTCGAGCCTGACGGA
CACTGGGTTCAGGCTGGCATCATCAGCTTTGCATCAAGCTGTGCCAGGAGGACGCTCCTGT
GCTGTGACCAACACAGCTGCTCACAGTTCTTGCTGCAGGCTCAGGTTTCAGGGGGCAGCTT
TCCTGGCCAGAGCCAGAGACCCCGAGATGAGTGATGAGGACAGCTGTGTAGCCTGTGGA
TCCTTGAGGACAGCAGGTCCCCAGGCAGGAGCACCTCCCCATGGCCCTGGGAGGCCAGGCT
GATGCACCAAGGACAGCTGGCCTGTGGCGGAGCCCTGGTGTGAGGAGGGCGGTGCTAACTG
CTGCCCCTGCTTCATTGGGCGCCAGGCCCCAGAGGAATGGAGCGTAGGGCTGGGGACCAGA
CCGGAGGAGTGGGGCCTGAAGCAGCTCATCCTGCATGGAGCCTACACCCACCTGAGGGGGG
CTACGACATGGCCCTCCTGCTGCTGGCCAGCCTGTGACACTGGGAGCCAGCCTGCCGCCCC
TCTGCCTGCCCTATCCTGACCACCACTGCTGATGAGGGAGCGTGGCTGGGTCTGGGACGG
GCCCGCCAGGAGCAGGCATCAGCTCCCTCCAGACAGTGCCCGTGACCCCTCCTGGGGCCTAG
GGCCTGCAGCCGGCTGCATGCAGCTCCTGGGGGTGATGGCAGCCCTATTCTGCCGGGATGG
TGTGTACCAAGTGTGTGGGTGAGCTGCCAGCTGTGAGGGCCTGTCTGGGGCACCCTGGTG
CATGAGGTGAGGGGCACATGGTTCTTGGCCGGGCTGCACAGCTTCGGAGATGCTTGCCAAGG
CCCCGCCAGGCCGGCGGTCTTACCGCGCTCCCTGCCTATGAGGACTGGGTGAGCAGTTTG
ACTGGCAGGTCTACTTCGCCGAGGAACAGAGCCCGAGGCTGAGCCTGGAAGCTGCCCTGGCC
AACATAAGCCAACCAACAGCTGCTGAGCAGGGGACCTGGCCATTCTCAGGACAAGAGAATGC
AGGCAGGCAAATGGCATTACTGCCCTGTCTCTCCCACTGTCTATGTGTGATTCAGGCAC
CAGGGCAGGCCAGAGGCCAGCAGCTGTGGGAAGGAACCTGCCTGGGGCCACAGGTGCCCA
CTCCCCACCTGCAGGACAGGGGTGTCTGTGGACACTCCACACCCAACCTGTCTACCAAGC
AGCGCTCTCAGCTTCTCTCTCTTACTCTTTCAGATACAATCACGCCAGCCACGTGTGTTT
TGAAAATTTCTTTTTTGGGGGGCAGCAGTTTTCTTTTTTAAACTTAAATAAATTGTTAC
AAAATAAAA

FIGURE 49

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA40571

MLLSSLVSLAGSVYLAWILFFVLVDFCIVCITTYAINVSLMWLSFRKVQEPQGKAKRHGNTV
PGEWPWQASVRRQGAHICSGSLVADTWVLTAAHCFEKAATELNSWSVVLGSLQREGLSPGA
EEVGVAALQLPRAYNHYSQGSDDLALLQLAHPHTHTPLCLPQPAHRFFPGASCWATGWDQDTS
DAPGTLRLNRLRLISRPTCNCIYNQLHQRHLSNPARPGMLCGGPQPGVQGPCQGDSSGGPVLC
LEPDGHWVQAGIISFASSCAQEDAPVLLTNTAAHSSWLQARVQGAFLAQSPETPEMSDEDS
CVACGSLRTAGPQAGAPSPWPWEARLMHQQLACGGALVSEEAVLTAAHCFIGRQAPPEWSV
GLGTRPEEWGLKQLILHGAYTHPEGGYDMALLLLAQPVTLGASLRPLCLPYPDHHLDPDGERG
WVLGRARPGAGISSLQTVFPVTLTGPRACSRLLHAAPGGDGSPIPLGMVCTSAVGELPSCEGLS
GAPLVHEVRGTWFLAGLHSFGDACQGPAPPAVFTALPAYEDWVSSLDWQVYFAEEPEPEAEP
GSCLANISQPTSC

Important features:

Signal peptide:

amino acids 1-15

Homologous region to Serine proteases, trypsin family

amino acids 79-95, 343-359 and 237-247

N-glycosylation sites.

amino acids 37-40 and 564-567

Kringle domains

amino acids 79-96, 343-360 and 235-247

FIGURE 50

CGGGCCGCCCCCGGCCCCCATTCGGGCCGGGCCTCGCTGCGGCGGCGACTGAGCCAGGCTGG
GCCGCGTCCCTGAGTCCCAGAGTCGGCGCGGCGCGCAGGGGCAGCCTTCCACCACGGGGAG
CCCAGCTGTACAGCCGCTCACAGGAAGATGCTGCGTCGGCGGGGCAGCCCTGGCATGGGTGT
GCATGTGGGTGCAGCCCTGGGAGCACTGTGGTTCTGCCTCACAGGAGCCCTGGAGGTCCAGG
TCCCTGAAGACCCAGTGGTGGCACTGGTGGGCACCGATGCCACCTGTGCTGCTCCTTCTCC
CCTGAGCCTGGCTTCAGCCTGGCACAGCTCAACCTCATCTGGCAGCTGACAGATACCAAACA
GCTGGTGCACAGCTTTGTGTAGGGCCAGGACCAGGGCAGCGCCTATGCCAACCGCACGGCCC
TCTTCCCGACCTGCTGGCACAGGGCAACGCATCCCTGAGGCTGCAGCGCGTGCCTGTGGCG
GACGAGGGCAGCTTCACCTGCTTCGTGAGCATCCGGGATTTTCGGCAGCGCTGCCGTGAGCCT
GCAGGTGGCCGCTCCCTACTCGAAGCCCAGCATGACCCTGGAGCCCAACAAGGACCTGCGGC
CAGGGGACACGGTGACCATCACGTGCTCCAGCTACCAGGGCTACCCTGAGGCTGAGGTGTTT
TGGCAGGATGGCAGGGTGTGCCCTGACTGGCAACGTGACCACGTGCGAGATGGCCAACGA
GCAGGGCTTGTTTGATGTGCACAGCGTCTGCGGGTGGTGCTGGGTGCGAATGGCACCTACA
GCTGCCTGGTGCACAACCCCGTGTGTCAGCAGGATGCGCACRGTCTGTCAACATCACAGGG
CAGCCTATGACATTCCCCCAGAGGCCCTGTGGGTGACCGTGGGGCTGTCTGTCTGTCTCAT
TGCACCTGCTGGTGGCCCTGGCTTTTCGTGTGCTGGAGAAAGATCAAACAGAGCTGTGAGGAG
AGAATGCAGGAGCTGAGGACCAGGATGGGGAGGGAGAAGGCTCCAAGACAGCCCTGCAGCCT
CTGAAACACTCTGACAGCAAAGAAGATGATGGACAAGAAATAGCCTGACCATGAGGACCAG
GAGCTGCTACCCCTCCCTACAGCTCCTACCCCTCTGGCTGCAATGGGGCTGCACTGTGAGCCC
TGCCCCCAACAGATGCATCCTGCTCTGACAGGTGGGCTCCTTCTCCAAGGATGCGATACAC
AGACCACTGTGCAGCCTTATTTCTCCAATGGACATGATTTCCAAAGTCATCCTGCTGCCTTTT
TTCTTTATAGACACAATGAACAGACCACCCACAACCTTAGTTCTCTAAGTCATCCTGCCTGCT
GCCTTATTTTACAGTACATACATTTCTTAGGGACACAGTACACTGACCACATCACCACCCCT
TTCTTCCAGTGTGCGTGGACCATCTGGCTGCCTTTTTTCTCCAAAGATGCAATATTTCAGA
CTGACTGACCCCTGCGCTTATTTACCAAAGACACGATGCATAGTCAACCCGGCCTTGTTTT
TCCAATGGCCGTGATACACTAGTGATCATGTTTACGCCCTGCTTCCACCTGCATAGAATCTTT
TCTTCTCAGACAGGGACAGTGCGGCCTCAACATCTCCTGGAGTCTAGAAGCTGTTTCTTTT
CCCTCCTTCTCCCTGCCCAAGTGAAGACAGGGCAGGGCCAGGAATGCTTTGGGGACACCG
AGGGGACTGCCCCCACCACCACATGGTGCTATTCTGGGCTGGGGCAGTCTTTTCTGCG
TTGCTCTGGCCAGCTCCTGGCCTCTGGTAGAGTGAGACTTCAGACGTTCTGATGCCTTCCG
GATGTCATCTCTCCTGCCCCAGGAATGGAAGATGTGAGGACTTCTAATTTAAATGTGGGAC
TCGGAGGGATTTGTAAACTGGGGGTATATTTGGGGAAAATAAATGTCTTTGTAAAAAAA
AAAAAAAAAAAAA

FIGURE 51

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA41386

><subunit 1 of 1, 316 aa, 1 stop, 1 unknown

><MW: -1, pI: 4.62, NX(S/T): 4

MLRRRGSPGMGVHVGAAALGALWFCLTGALEVQVPEDPVVALVGTDATLCCSFSPGPGFSLAQ
LNLIWQLTDTKQLVHSFAEGQDQGSAYANRTALFPDLLAQGNASLRLQVRVVADEGSFTCFV
SIRDFGSAAVSLQVAAPYSKPSMTLEPNKDLRPGDTVTITC SSYQGYPEAEVFWQDGGQGVPL
TGNVTT SQMANEQGLFDVHSLRVVLGANGTYSCLVRNPVLQQDAHXSVTITGQPMTFPPEA
LWVTVGLSVCLIALLLVALAFVCWRKIKQSC EEENAGAEDQDGE GEGSKTALQPLKHSDSKED
DGQEIA

Important features:

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 251-270

N-glycosylation site.

amino acids 91-94, 104-107, 189-192 and 215-218

Homologous region to Immunoglobulins and MHC

amino acids 217-234

FIGURE 52

TTCTGTACCCCTTGAGAAAAGAGTTGGTGGTAAATGTGCCACGTCTTCTAAGAAGGGGGAGTC
CTGAACCTTGTCTGAAGCCCTTGTCCGTAAGCCTTGAACCTACGTTCTTAAATCTATGAAGTCG
AGGGACCTTTTCGTGCTTTTGTAGGGACTTCTTTCCTTGCTTCAGCAACATGAGGCTTTTCT
TGTGGAACCGCGTCTTGACTCTGTTTCGTCACTTCTTTGATTGGGGCTTTGATCCCTGAACCA
GAAGTGAAAATTGAAGTTCTCCAGAAGCCATTATCTGCCATCGCAAGACCAAAGGAGGGGA
TTTGATGTTGGTCCACTATGAAGGCTACTTAGAAAAGGACGGCTCCTTATTTCACTCCACTC
ACAAACATAACAATGGT CAGCCCATTTGGTTTACCCTGGGCATCCTGGAGGCTCTCAAAGGT
TGGGACCAGGGCTTGAAAGGAATGTGTGTAGGAGAGAAGAGAAAGCTCATCATTCCTCTCTGC
TCTGGGCTATGGAAAAGAAGGAAAAGGTAATAATCCCCAGAAAGTACACTGATATTTAATA
TTGATCTCCTGGAGATTGAAATGGACCAAGATCCCATGAATCATTCCAAGAAATGGATCTT
AATGATGACTGGAACCTCTCTAAGATGAGGTTAAAGCATATTTAAAGAAGGAGTTTGAAAA
ACATGGTGCGGTGGTGAATGAAAGTCATCATGATGCTTTGGTGGAGGATATTTTGATAAAG
AAGATGAAGACAAAGATGGGTTTATATCTGCCAGAGAATTTACATATAAACACGATGAGTTA
TAGAGATACATCTACCCCTTTTAAATATAGCACTCATCTTTCAGAGAGGGGAGTCATCTTTAA
AGAACCTTTTATTTTATACAATGTTCTTTCTTGCTTTGTTTTTTATTTTATATATTTTGT
CTGACTCTTATTTAAAGAACCCCTTAGGTTTCTAAGTACCCATTTCTTTCTGATAAGTTATT
GGGAAGAAAAGCTAATTGGTCTTTGAATAGAAGACTTCTGGACAATTTTTCACCTTTACAG
ATATGAAGCTTGTTTTACTTTCTCACTTATAAATTTAAATGTTGCAACTGGGAATATACC
ACGACATGAGACCAGGTTATAGCACAAATTAGCACCTATATTTCTGCTTCCCTCTATTTTC
TCCAAGTTAGAGGTCAACATTTGAAAAGCCTTTTGCAATAGCCCAAGGCTTGCTATTTTCAT
GTTATAATGAAATAGTTTATGTGTAAGTGGCTCTGAGTCTCTGCTTGAGGACCAGAGGAAAA
TGGTTGTTGGACCTGACTTGTGTTAATGGCTACTGCTTTACTAAGGAGATGTGCAATGCTGAAG
TTAGAAAACAAGTTAATAGCCAGGCATGGTGGCTCATGCCGTGAATCCCAGCACTTTGGGAG
GCTGAGGCGGGCGGATCACCTGAGGTTGGGAGTTCGAGACCAGCCTGACCAACACGGAGAAA
CCCTATCTCTACTAAAAATACAAGTAGCCCGCGTGGTGATGCGCTGCCTGTAATCCCAGCT
ACCCAGGAAGGCTGAGGCGGCAGAAATCACTTGAACCCGAGGCCGAGGTTGCGGTAAGCCGAG
ATCACCTNCAGCCTGGACACTCTGTCTCGAAAAAAGAAAAGAACACGGTTAATACCATATNA
ATATGTATGCATTGAGACATGCTACCTAGGACTTAAGCTGATGAAGCTTGGCTCCTAGTGAT
TGGTGGCCTATTATGATAAATAGGACAAATCATTATGTGTGAGTTTCTTTGTAATAAAATG
TATCAATATGTTATAGATGAGGTAGAAAGTTATATTTATATTTCAATATTTACTTCTTAAGGC
TAGCGGAATATCCTTCTGGTCTTTAATGGGTAGTCTATAGTATATTTACTACAATAACA
TTGTATCATAGATAAAGTAGTAAACCAGTCTACATTTTCCATTTCTGTCTCATCAAAAAC
TGAAGTTAGCTGGGTGTGGTGGCTCATGCCGTGAATCCCAGCACTTTGGGGGCCAAGGAGGG
TGGATCACTTGAGATCAGGAGTTCAAGACCAGCCTGGCCAACATGGTGAAACCTTGTCTCTA
CTAAAAATACAAAAATTAGCCAGGCGTGGTGGTGACACCTGTAGTCCAGCTACTCGGGAG
GCTGAGACAGGAGATTGTCTTGAACCCGGGAGGCGGAGGTTGCAGTGAGCCAAGATTGTGCC
ACTGCACTCCAGCCTGGGTGACAGAGCAAGACTCCATCTCAAAAAAAGAAAGCAGAG
CCTACAGCAGCTACTATTGAATAAATACCTATCTCGGATTTT

00573403-111611

090310

><subunit 1 of 1, 211 aa, 1 stop

MRLFLWNVAVLTLFVTSLIGALIPPEVVKIEVLQKPFICHRKTKGGDLMLVHYEGYLEKDGSL
 FHSTHKHNNGQPIWFTLGLILEALKGWDQGLKGMCVGEKRKLIIPPALGYGKEGKGKIPPEST
 LIFNIDLLLEIRNGPRSHESFQEMDNLDDWKLKDEVKAYLKKFEFEKHGAVVNESHHDALVED
 IFDKEDDKDQGFISAREFTYKHDEL

Signal peptide:

amino acids 1-20

amino acids 176-179

amino acids 143-146, 156-159, 178-181 and 200-203

amino acids 208-211

amino acids 78-114 and 118-131

amino acids 191-203, 184-203 and 140-159

amino acids 183-203

FIGURE 54

AATAAAGCTTCCTTAATGTTGTATATGTCTTTGAAGTACATCCGTGCATTTTATTTTATAGCAT
CCAACCATTCCTCCCTTTGTAGTTCTCGCCCCCTCAAATCACCCCTCTCCCGTAGCCCCACCGA
CTAACATCTCAGTCTCTGAAAATGACACAGAGATGCCTGGCTACCTCGCCCTGCCCTTCAGCCT
CACGGGGCTCAGTCTCTTTTCTCTTTGGTGCCACCAGGACGGAGCATGGAGGTCACAGTAC
CTGCCACCCCTCAACGTCTCAATGGCTCTGACGCCCGCTGCCCTGCACCTTCAACTCCTGC
TACACAGTGAACCACAAACAGTTCTCCCTGAACTGGACTTACCAGGAGTGCAACAACTGCTC
TGAGGAGATGTTCTCCTCCAGTTCGCATGAAGATCATTAACTGAACTGGAGCGGTTTCAAG
ACCGCGTGGAGTCTCAGGGAACCCAGCAAGTACGATGTGTCCGGTGATGCTGAGAAACGTG
CAGCCGGAGGATGAGGGGATTTACAACTGCTACATCATGAACCCCTGACCGCCACCGTGG
CCATGGCAAGATCCATCTGCAGGTCTCATGGAAGAGCCCCCTGAGCGGGACTCCACGGTGG
CCGTGATTGTGGGTGCCCTCCGTGGGGGCTTCTCGGCTGTGGTTCATCTTGGTGCTGATGGTG
GTCAAGTGTGTGAGGAGAAAAAAGAGCAGAAGCTGAGCACAGATGACCTGAAGACCGAGGA
GGAGGGCAAGACGGACGGTGAAGGCAACCCGGATGATGGCGCCAAGTATAGTGGGTGGCCGGCC
CTGCAGCCTCCCGTGTCCCGTCTCCTCCCTCTCCGCCCTGTACAGTGACCTGCCTGCTCG
CTCTTGGTGTGCTTCCCGTGACCTAGGACCCAGGGCCCACTGGGGCTCCTGAACCCCCG
ACTTCGTATCTCCACCTGACCAAGAGTGACCCACTCTCTTCCATCCGAGAAACCTGCCA
TGCTCTGGGACGTGTGGGCCCTGGGGAGAGGAGAGAAAGGCTCCACCTGCCAGTCCCTGG
GGGGAGGCAGGAGGCACATGTGAGGGTCCCAGAGAGAAGGGAGTGGGTGGGCAGGGGTAGA
GGAGGGGCGCTGTCACTGCCAGATGCTTGCTGGCAGTGGCTTACAGAGAGACCTGGTGG
GGAGGGAGGGCTTTCCTGTGCTGACAGCGCTCCCTCAGGAGGGCTTGGCTTGGCAGGGCTG
TGCTCCTCCCTGCTCCAGCCAGAGCAGCCATCAGGCTGGAGGTGACGATGAGTTCCTGA
AACTTGGAGGGGCATGTTAAAGGGATGACTGTGCATTCCAGGGCACTGACGGAAGCCAGGG
CTGCAGGCAAAGCTGGACATGTGCCCTGGCCAGGAGGCCATGTTGGGCCCTCGTTTCCATT
GCTAGTGGCCTCCTTGGGGCTCCTGTTGGCTCCTAATCCCTTAGGACTGTGGATGAGGCCAG
ACTGGAAGAGCAGCTCAGGTAGGGGGCCATGTTTCCAGCGGGGACCCACCAACAGAGGCC
AGTTTCAAAGTCACTGAGGGGTGAGGGGTGGGGCTCCATGGTGAATGCAGGTTGCTGCAG
GCTCTGCCTTCTCCATGGGGTAACACCTCGCCTGGGCAGGGGCAGCCAAGGCTGGGAAAT
GAGGAGGCCATGCACAGGGTGGGGCAGCTTCTTTGGGGCTTCACTGAGAACTCTCCAGTT
GCCCTTGGTGGGGTTTCCACCTGGCTTTTGGCTACAGAGAGGGGAAGGGAAAGCCTGAGGCCG
GCATAAGGGGAGGCCCTTGAACCTGAGCTGCCAATGCCAGCCCTGTCCCATCTGCGGCCACG
CTACTCGCTCCTCTCCCAACAACTCCCTTCGTGGGGACAAAAGTGACAATTGTAGGCCAGGC
ACAGTGGCTCACGCTGTAATCCCAGCACTTTGGGAGGCCAAGGCGGGTGGATTACCTCCAT
CTGTTTAGTAGAAATGGGCAAAACCCCATCTCTACTAAAAATACAAGAATTAGCTGGGCGTG
GTGGCGTGTGCTGTAATCCCAGCTATTGGGAGGCTGAGGCAGGAGAATCGCTTGAAGCCG
GGAAGCAGAGGTTGCAGTGAACCTGAGATAGTAGTGCCACTGCAATTACGCTGGGTGAC
ATAGAGAGACTCCATCTCAAAAAA

FIGURE 55

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45415

<subunit 1 of 1, 215 aa, 1 stop

<MW: 24326, pI: 6.32, NX(S/T): 4

MHRDAWLPRPAFSLTGLSLFFSLVPPGRSMEVTVPATLNLVNGSDARLPCTFNSCYTVNHKQ
FSLNWTYQECNNCSEEMFLQFRMKIINLKLERFQDRVEFSGNPSKYDVSVMRLRVQPEDGEI
YNCYIMNPDRHRGHGKIHLQVLMEEPPERDSTVAVIVGASVGGFLAVVILVLMVVKCVRRK
KEQKLSTDDLKTEEEGKTDGEGNPDDGAK

Important features:

Signal peptide:

amino acids 1-20

Transmembrane domain:

amino acids 161-179

Immunoglobulin-like fold:

amino acids 83-127

N-glycosylation sites.

amino acids 42-45, 66-69 and 74-77

GTTGTATATGTCCTGAAGTACATCCGTGCATTTTTTTTTAGCATCCAACCATCCTCCCTTGTA
 GTTCTCGCCCCCTCAAATCACCTTCTCCCTTAGCCCACCNACTAACATCTCAGTCTCTGAA
 AATGCACAGAGATGCCTGGCTACCTCGCCCTGCCTTCAGCCTCACGGGGCTCAGTCTCTTTT
 TCTCTTTGGTGCCACCAGGACGGAGCATGGAGGTCACAGTACCTGNCCACCCTCAACGTCC
 TCAATGGCTCTGACGCCCCGCTGCCCTGCCCTTCAACTCCTGCTACACAGTGAACCACAAAC
 AGTTCTCCCTGAACTGGACTTACCAGGAGTGCAACAAC TGCTCTGAGGAGATGTTCTCCAG
 TTCGCGATGAAGATCATTAACCTGAAGCTGGAGCGGTTTCAAGACCGCGTGGAGTTCTCAGG
 GAACCCACGCAAGTACGATGTGTCGGTGATGCTGAGAAACGTGCAGCCGAGGATGAGGGGA
 TTTACAAC TGCTACATCATGAACCCCCC

FIGURE 57

TCACGGGGCTCATCTCTTTTTCTCTTTGGTGCCCACCAGGACGGAGCATGGAGGTNCACATA
CCTGCCACCCCTCAACGTCCCTCAATGGCTTTGACGCCCGCCTGCCCTGCACCTTCAACTCCNG
CTACACAGTGAACCACAAACAGTTCTCCCTGAACTGGATTTACCAGGAGTGCAACAACCTGGC
TCTGAGGAGATGTTCTCTCAGTTCCCGCATGGAAGATCATTTAACCTGAAAGCTGGAAGCGG
TTTTCAGAACC CGTGGAAGTTTCTCAGGGAACCCAGCAAGTACGATGTGT CGGTGATGC
TGAGAAACGTGCAGCCGGAGGATGAGGGGATTTACAAC TGCTACATCATGAACCCCCC

Sequence

[illegible]

TGCGGCGACCGCTCGTACACATGCGGCTCCACCTTCGCCCTTACCGTGTGGGGCTGCTCCCGGATGGCTCTCTGT
TCCTCTTCTGCTCTGCTAATGCTGCTCGGGACCCAGCGCTCCCGGCCGACGCTACACCCCTAGTGTGCTGTGCTC
CTGGTGAATTTGGGTAAACCAACTGGAGCAAGCTGAGCAAGCCAGAGTGGTGCACTACCTCTGCTCCAGAAAG
CCGAAAGCTACTTACAACTCTGCTGAACTGGAAGCTGCTGCTGCTGTGTCATCATCTGACTGCTGGATTGACAATA
TCAGGCTGTGTTTACACAAACATCAGGGCCACCCAGTTCTGTATGGTGTGGATGTAGCTGTCTCCCTGGCTTTG
GGAAGCACTTCTCACTGGAGTCTCTGGACCCAGCAAAAGCAGCGTGGGTCTCTATTTCACACCATGTTGGAGA
GCCTTGTGGGCTGGGGCTACACACGGGGTGGAGTGTCTCGAGGGCTCCCTTACTGACTGGCCGGAGCCCAAAT
AAAACGGGCCCTACTTCTGCGCTCCGGAGATGATCGAGGAGATGTACAGCTGTATGTTGGGGCCCCGTGGTG
TGATGCGCCACAGTATGGGCAACATGTACACGCCTCATTTCTCGAGCGCGAGCCGAGGCTGGGAAGCAAGT
TGTCTCGGGCTCTGTGCTGCTACGTGGTGCCTCTGGGGGGCTGGCCAGAACCTTGGCGTCTGGTCTCAGAG
ACAAACAACCGGATCCCGATCTAGGGCCCTGAAGATCGGAGCAGCGCTGAGCTGTCTCCACACAGCTGGG
TGCTGTCCCTTACAACATACATGTCACCTGTGAGAAGGTTGTTGGCAGACACCACTAACTACACAGCTGCGGG
ACTACCGCAAGTCTTTCAGGACATCGGCTTTGAAGATGGCTGGCTCATCGCGCAGGACACAGAAGGGCTGTGTG
AAGCCACAGTGGCCACTTGGGCTGCAGCTGCACCTGCTATGTGTAATGGCTGCCACACAGACTCTTCTTACT
ATAGAGATCTTCCCTGACCTGACCTAAATCTGCTTTGGTGACGGCAGTGTGATGTGAATCTGAAGAGTGGCC
TGCAGTGCGAGGCTGGCAGACGCTGCCAGGAGCACAAGTGTGTGCTCGAGGACTCGCAGGACGAGACATCT
AGATGTGCGCCACGACGACCACTCGGCCATCTGAACAACTGTGCTCTTGGGCCCTGCACTCTGTGACACAGGA
CTCCTGTGGCTCGGCCCTGCACTGTGCTTGGCCTCTGGGGCTCATGGGCAACCGCTTTTGGCAAGTGTGTGCT
CTCACATTCAAGGCCCGAGTCTTGGAGCTGTGAAGCATCTGCCATGGGGAAGTGTGTTGTATCTCTTCTTCT
GTGGCAGTGAAGGAAGGAATGAGAGTCTAGACTCAAGGGACATGGATGGCAAGATGCTGCTGATGGTGA
ACTGTGCTGACACTTAGGACTGGCTCTACAGGGTGGAGCTGTGGCCCTGGTCTCCAGTCTGCTGCTGGGCCAT
TGTCCCTCTATTCTGTGGGCTTTTATACTTGCTACTTGGCCCTGGCCCGCAGCTCTCTATGAGGATGTT
ACTGGGCTGTGCTCTGTATCCGAGAGTCCAGGAGATGGCTCTGCGCCCTGGGTGAACCTTCCCAACACACA
GCCACAGATGAGGCTTGCACATGGATCTGGGTAGCTAGAGCTGTGGTCTCTCTGGCTAGCTGTGGCCAGCG
TGACTGGCTCTCTGGGCGAGCCTAGTAGCTCTTGCAGGCGAGGACATTTAGCTCTTCTGTGTTTCCAGGAG
CTTGGGACATCTCACTCACTCTACTCTCCCTTACCAACAGGAGCATTTCAAGCTCTGGATGGGCAGCATGTG
CCCCAGTCCCGCAGGCTGTGTTCTCAGGGGCCGATTTCCTGGATGTGCTATTGGCCCCAGGATGAAGCTGC
CTCCTTCTACCTCTGGGACTGTGGTTCTCAAGGATGAGACAGGGGTGGAGCGCATGCCCCTTGGGAACCTTAGGA
GAAAGGAATTCAGGAAGCAGCCAAAGCTGTCTCGAGCTTCCCTGAGTGCATCTCTTGCTAACCCACCATCA
CACTGCCACCTCTAGGCTTAGGCTCTCACTAGTACCAAGTGGGTCGACGACCGGCTGAGGATGGGGCTCTTATCC
CTGGCCAGCACCAGCTCTGTTGTCTGGAGTAGGCCAGTCCAGAAAGTTGATAGGAGCCCTGAGAGACGAGGGTCCC
TGAGGCCCCCTAGGGGCTTCTGTCTGCCCCAGGGTGTCTCATGTGATCTCCTGTGGGACGAGCATGAGAGT
CAGGGCTGCTCTTTCAGCGTAGGCTCTAAGTGGTGTGCTAGGCCAGCCGAGAAAGGTACAGCTCTAGTGT
GGGTCTCCCAAGGACGCTTCAGCTGGAATGAGCTGCTCTCCACAGGGTTTCTGTGACAGTGGATTTTCTCTG
TGTACATACCTGCTGGCATCTGTCTCCCTTGTCTCTGATGGGCCCATCTGGGCTCTGAGCAGGCTGTATCTG
GATTCTGGCAATTAAGAGTACTCTGATGTGTTAAAAAATAAAAAAATAAAAAA

FIGURE 59

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA44189

><subunit 1 of 1, 412 aa, 1 stop

><MW: 46658, pI: 6.65, NX(S/T): 4

MGLHLRPFYRVGLLEDGLFLLLLLLMLLADPALPAGRHPPVVLVPGDLGNQLEAKLLDKPTVVH
YLCSKKTESYFTIWLNLLELLLPVIIDCWIDNIRLVYNKTSRATQFPDGDVVRVPGFGKTFSL
EFLDPSKSSVGSYFHTMVESLVGWGYTRGEDVRGAPYDWRRAPNENGPYFLALREMIEMYQ
LYGGPVLVAHSMGNMYTLYFLQRQPQAWKDKYIRAFVSLGAPWGGVAKTLRVLASGDNNRI
PVIGPLKIREQQRSVSTSWLLPYNYTWSPEKVFTPTINYTLRDYRKFFQDIGFEDGWL
RQDTEGLVEATMPPGVQLHCLYGTGVPTPDSFYYESFPDRDPKICFGDGDGTVNLKSALQCQ
AWQSRQEHQVLLQELPGSEHIEMLANATTLAYLKRVLG

Important features:

Signal peptide:

amino acids 1-28

Potential lipid substrate binding site:

amino acids 147-164

N-glycosylation sites.

amino acids 99-102, 273-276, 289-292 and 398-401

Lipases, serine proteins

amino acids 189-201

Beta-transducin family Trp-Asp repeat

amino acids 353-365

CGGACGCGTGGGCGGACGCGTGGGGCGGCGGACGGCGGGCAGCGGCACATGGAGAGCGGG
GCCTACGGCGCGGCCAAGGCGGGCGGCTCCTTCGACCTGCGGCGCTTCCTGACGACGCCGA
GGTGTGTGGCGCGCGCCGTGTGCTTGGTCTTCGCCTTGATCGTGTCTCCTGCATCTATGGTG
AGGGGTACAGCAATGCCACGAGTCTAAGCAGATGTACTGCGTGTTC AACCGCAACGAGGAT
GCCTGCCGCTATGGCAGTGCCATCGGGGTGCTGGCCTTCCTGGCTCGGCCTTCTTCTTGG
GGTCGACGCGTATTTCCCCCAGATCAGCAACGCCACTGACCGCAAGTACCTGGTCATTGGTG
ACCTGCTCTTCTCAGCTCTCTGGACCTTCTGTGTTTGTGTGGTTTCTGCTTCTCACC AAC
CAGTGGGCAGTCACCAACCCGAAGGACGTGCTGGTGGGGGCCGACTCTGTGAGGGCAGCCAT
CACCTTCAGCTTCTTTTCCATCTTCTCCTGGGGTGTGCTGGCCTCCCTGGCCTACCAGCGCT
ACAAGGCTGGCGTGGACGACTTCATCCAGAATTACGTTGACCCCCACTCCGAGCCCCAACACT
GCCTACGCCCTCTACCCAGGTGCATCTGTGGACAAC TACCAACAGCCACCCTTACCCAGAA
CGCGGAGACCACCGAGGGCTACCAGCCGCCCCCTGTGTACTCGATGGCGGTTAGCGTGGGAA
GGGGGACAGAGAGGGCCCTCCCCTCTGCCCTGGACTTTCCCATCAGCCTCCTGGAACTGCCA
GCCCTCTCTTTACCTGTTCCATCCTGTGCAGCTGACACACAGCTAAGGAGCCTCATAGCC
TGGCGGGGGCTGGCAGAGCCACACCCCAAGTGCCTGTGCCAGAGGGCTTCAGTCAGCCGCT
CACTCCTCCAGGGCACTTTTAGGAAAGGGTTTTTAGCTAGTGTTTTTCTCGCTTTTAATGA
CCTCAGCCCCGCTGCA GTGGCTAGAAGCCAGCAGGTGCCCATGTGCTACTGACAAGTGCCT
CAGCTTCCCCCGGCCCGGGTFCAGGCCGTGGGAGCCGCTATTATCTGCGTTCTCTGCCAAAG
ACTCGTGGGGCCATCACACTGCCTGTGCAGCGGAGCCGACCAGGCTCTGTGTCTCTCA
CTCAGTTTGCTTCCCTGTGCCCACTGCTGTATGATCTGGGGGCCACCACCTGTGCCGGT
GGCCTCTGGGCTGCCTCCCGTGGTGTGAGGGCGGGGCTGGTGTCTATGGCACTTCTCCTTG
CTCCACCCCTGGCAGCAGGGAAGGGCTTTGCCTGACAACACCCAGCTTTATGTAATAATTC
TGCA GTTGTTACTTAGGAAGCCTGGGGAGGGCAGGGGTGCCCCATGGCTCCCAGACTCTGTCT
TGTGCCGAGTGTATTATAAAATCGTGGGGGAGATGCCCGGCTGGGATGCTGTTTGGAGACG
GAATAAATGTTTTCTCATTCAAAG

<subunit 1 of 1, 224 aa, 1 stop

MESGAYGAAGGSFDLRRFLTPQPVVARAVCLVFALIVFSCIYEGEGYSNAHESKQMYCVFN
RNEDACRYGSAIGVLAFLASAFFLVVDAYFPQISNATDRKYLIVIGDLLFSALWTLWFVGF
FLTNQWAVTNPKDVLVGADSVRAITFSFFSIFSWGVLASLAYQRYKAGVDDFIQNYVDPTP
DPNTAYASYPGASVDNYQOPPPTQNAETTEGYOPPPVY

Type II Transmembrane domain:

amino acids 1-45

amino acids 74-90, 108-126 and 145-161

amino acids 97-100

FIGURE 63

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA49152
><subunit 1 of 1, 802 aa, 1 stop
><MW: 88846, pI: 6.41, NX(S/T): 7
MPVAEAPQVAGGQGDGGDGEEAEPEGMFKACEDSKRKARGYLRLVPLFVLLALLVLASAGVL
LWYFLGYKAEVMVSQVYSGSLRVLNRHFSQDLTRRESSAFRSETAKAQKMLKELITSTRLGT
YNNSSSVYSFGEGLTCTFFWFILQIPEHRRLMLSPEVVQALLVEELLSTVNSSAAVPYRAEY
EVDPEGLVILEASVKDIAALNSTLGCYRYSYVQGQVLRLKGPDLASSCLWHLQGPKDMLM
KLRLEWTLAECRDLAMYDVAGPLEKRLITSVYGCSRQEPVVEVLASGAIMAVVWKKGLHSY
YDPFVLSVQPVVFQACEVNLTLNRLDSQGVLTSTPYFPSYSPQTHCSWHLTVPSLDYGLAL
WFDAYALRRQKYDLPTQGQWTIQNRRLCGLRILQPYAERIPVVATAGITINFTSQISLTGP
GVRVHYGLYNQSDPCPGEFLCSVNGLCVPACDGVKDCPNGLDERNCVCRATFQCKEDSTCIS
LPKVCDCQPDCLNGSDEEQCQEGVPCGTFTFQCEDRSCVKKPNPQCDGRPD CRDGSDEEHCD
CGLQGPPSRIVGGAVSSEGEWPWQASLQVRGRHICG GALIADRWVITAAHCFQEDSMAS TVL
WTVFLGKVVQNSRWPGEVSKVSRLLLLHPYHEEDSHDYDVALLQLDHPVVRSAAVRPVCLPA
RSHFFEPGLHCWITGWGALREGGPISNALQKVDVQLIPQDL CSEAYRYQVTPRMLCAGYRKG
KKDACQDGGSGPLVCKALSGRWFLAGLVSWGLGCGRPNYFGVYTRITGVISWIQQVVT
```

Important features:

Type II transmembrane domain:

amino acids 46-67

Serine proteases, trypsin family, histidine active site.

amino acids 604-609

N-glycosylation sites.

amino acids 127-130, 175-178, 207-210, 329-332, 424-427, 444-447
and 509-512

Kringle domains.

amino acids 746-758 and 592-609

Homologous region to Kallikrein Light Chain:

amino acids 568-779

Homologous region to Low-density lipoprotein receptor:

amino acids 451-567

FIGURE 64

GCACCCAGGGCCAGTGGACGATCCAGAACAGGAGGCTGTGTGGCTTGC GCATCCTGCAGCCC
TACGCCGAGAGGATCCCCGTGGTGGGCCACGGCCGGGATCACCATCAACTTCACCTCCCAGAT
CTCCCTCACCGGGCCCGGTGTGCGGGTGCACTATGGCTTGTACAACAGTCGGACCCCTGCC
CTGGAGAGTTCTCTGTCTGTGAATGGACTCTGTGTCCCTGCCCTGTGATGGGGTCAAGGAC
TGCCCCAACGGCTTGGATGAGAGAACTGCGTTTGCAGAGCCACATTCCAGTGCAAAGAGGA
CAGCATATGCATCTCACTGCCCAAGGTCTGTGATGGGCAGCCTGATTGTCTCAACGGCAGCG
ATGAAGAGCAGTGCCAGGAAGGGGTGCCATGTGGGACATTCACCTTCCAGTGTGAGGACCGG
AGCTGCGTGAAGAAGCCCAACCCGAGTGTGATGGGCGGCCGACTGCAGGGACGGCTCGGA
TGAGGAGCACTGTGACTGTGGCCTCCAGGGCCCTCCAGCCGCATTGTTGGTGGAGCTGTGT
CCTCCGAGGGTGAGTGGCCATGGCAGGCCAGCCTCCAGGTTGCGGGTTCGACACATCTGTGGG
GGGGCCCTCATCGCTGACCGCTGGGTGATAACAGCTGCCCACTGCTTCCAGGAGGACAGCAT
GGCCTCCACGGTGCTGTGGACCGTGTTCCTGGGCAAGGTGTGGCAGAACTCGCGCTGGCCTG
GAGAGGTGTCTTCAAGGTGAGCCGCTGCTCCTGCACCCGTACCACGAAGAGGACAGCCAT
GACTACGACGTGGCGCTGCTGCAGCTCGACCACCCGGTGGTGCCTCGGCCGCCGTGCGCCC
CGTCTGCCTGCCCGCGCGCTCCCACTTCTTCGAGCCCGGCCTGCACTGCTGGATTACGGGCT
GGGGCGCCTTGCGCGAGGGCGGGCCCATCAGCAACGCTCTGCAGAAAGTGGATGTGCAGTTG
ATCCCACAGGACCTGTGCAGCGAGGCCATATCGCTACCAGGTGACGCCACGCATGCTGTGTGC
CGGCTACCGCAAGGGCAAGAAGGATGCCGTGTCAGGGTGACTCAGGTGGTCCGCTGGTGTGCA
AGGCACTCAGTGGCCGCTGGTTCTTGGCGGGGCTGGTCAGCTGGGGCCTGGGCTGTGGCCGG
CCTAACTACTTCGGCGTCTACACCCGCATCACAGGTGTGATCAGCTGGATCCAGCAAGTGGT
GACCTGAGGAAGTGCCCCCTGCAAAGCAGGGCCACCTCCTGGA CTGAGAGCCCAGGGC
AAGTGCCAAGCAGGGGGACAAGTAT

[illegible]

FIGURE 66

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA49646

><subunit 1 of 1, 354 aa, 1 stop

><MW: 39362, pI: 8.35, NX(S/T): 2

MSNSVPLLCSFWSLCYCFAAGSPVPFGPEGRLEDKLHKPKATQTEVKPSVRFNLRSTSKDPEHE
GCYLSVGHSQPLEDCSFNMTAKTFFIIHGWTMSGIFENWLHKLVSALHTREKDANVVVDWL
PLAHQLYTDVANNTRVVVGHSIARMLDWLQEKDDFSLGNVHLIGYSLGAHVAGYAGNFVKGTV
GRITGLDPAGPMFEGADIIHKRLSPDDADFVDVLHTYTRSFGLSIGIQMPVGHIDIYPNGGDF
QPGCGLNDVLGSIAYGTITEVVKCEHERAVHLFVDSLVDKPSFAFQCTDSNRFKKGICLS
CRKNRCNSIGYNAKKMRNKRNSKMYLKTRAGMPFRGNLQSLECP

Important features:

Signal peptide:

amino acids 1-16

Lipases, serine active site.

amino acids 163-172

N-glycosylation sites.

amino acids 80-83 and 136-139

FIGURE 67

CGGACGCGTGGGCGGACGCGGTGGGCTGGGCAAGGCGCGGGCGCGGGCGGAGCCACCTCTTCCCTCCCCCGC
TTCCTGTGCGCTCCGCTGGCTGGACGCGCTGGAGGAGTGGAGCAGCACC CGGCGGCCCTGGGGGCTGACAGT
CGGCAAGATTGGGCCGAAGAGGAAGTGGTCTCAAACCCCGCAGGTGGCGACAGGCCAGACCAGGGGCGCTCG
CTGCCTGAGGGGCGGGCTGTAGGCGAGGGCGCGGCCAGTGCCGAGACCCGGGGCTTCAGGAGCGGGCCCCGGGAG
AGAAGCTGCGGGCGGGCGGAGCGGAGAAACAACTCCAAAGTTGGCGAAAGGCACCGCCCTACTCCCGGGCTGCCG
CGCGCTCCCCGCCCCAGCCCTGGCATCCAGAGTACGGGTGAGCGCCGAGCCCATGGAGCCCGCTGGGCGAGCGG
CACCAGGGAGCTGGGCGCCCGGGCTCCGCCGACCCCATCGGTAGACCACAGAAGCTCCGGGACCTTCCG
GCACCTCTGGACAGCCAGGATGCTGTGGCCACCTCTCTCTCTCTCTCTTGGAGGGGCTCTGGCCCATCCAG
ACCGGATTATTTTCCAAATCATGCTTGTGAGGACCCCCAGCAGTGTCTTAGAAGTGCAAGGCCACTTACAGA
GGCCCTTGGTCCGGGACAGCCGACCTTCCCTGGCCAACTGCACCTGGCTCATCTGGGCAGCAGGAACAGACTG
TCACCATCAGGTTCCAGAAGCTACACCTGGCCGTGTGGCTCAGAGCGCTTAACCCCTACGCTCCCTCTCCAGCCAC
TGATCTCCCTGTGTGAGGCACCTCCAGACCTCTGCGAGCTGCCCGGGGCAACGTACCATTCACTTACAGCTATG
CTGGGGCCAGAGCACCATGGGCGAGGGCTTCTGTCTCTCTACAGCCAAGATTGGCTGATGTGCCCTGCAAGAA
AGTTTCAGTGCTGAACCA CCGCTGTGTATCTGTCTGTCCAGCGTGTGATGGGGTTGATGCGCTGTGGCGATGGCT
CTGATGAAGCAGGTTGACGCTCAGACCCCTTCCCTGGCTGACCCCAAGACCCGCTCCCTCCCTTGGCTTGCAGT
TCACCTTGGAGGACTTCTATGGGGTCTTCTCTCTCTGGATATACACACTAGCCCTCAGTCTCCCAAGCGGAGT
CCTGCCATTGGCTGTGGACCCCATGATGGCCGGCGGCTGGCCGTGCGCTTACAGCCCTGGACTTGGGCTTTG
GAGATGTGAGTGTGTATGACGGCCCTGGGCCCCCTGAGAGCTCCCGACTACTGCGTAGTCTCACCCACTTCA
GCAATGGCAAGGCTGTCACTGTGGAGACACTGTCTGGCCAGGCTTGTGTCTTACCACACAGTTGCTTGGAGCA
ATGGTCTGTGGCTTCAATGCCACCTACCATGTGCGGGGCTATTGTCTGCTTGGGACAGACCCCTGTGGCTTAGGCT
CTGGGCTGGGAGCTGGGGAAGGCTAGGTGAGCGCTGCTACAGTGAGGCAAGCGCTGTGA CCGCTCATGGGACT
GTGCTGACGGCACAGATGAGGAGGACTTGCCAGGCTGCCCACTGGACACTTCCCTGTGGGGCTGCTGGCACCT
CTGGTGCCACAGCCTGTACTGTGCTGCTGACCGCTGCAACTACCAGACTTCTGTGCTGATGGAGCAGATGAGA
GACGCTGTGCGCATTGCCAGCCTGGCAATTTCCGATGCCGGGACGAGAAGTGCCTGTATGAGACGTGGGTGTGCG
ATGGGCGAGCCAGACTGTGCGGACGGCAGTGATGAGTGGGACTGCTCCTATGTTCTGCCCCGCAAGTCTATTACAG
CTGCGAGTCATTGGGAGCCTAGTGTGCGGCCGTCTCCTGGTTCATCGCCCTGGGCTGACCTGCAAGCTCTATGCCA
TTCGCAACCCAGGAGTACAGCATTTTCCCCCTCTCTCCCGATGGAGGCTGAGATTGTGCGAGCAGCAGGCACCCC
CTTCTACGGGCGAGCTCATTTGCCAGGGTCCATCCACCTGTAGAAGACTTTCCTACAGAGAATCTTAATGATA
ACTCAGTGCTGGGCAACCTGCGTTCTCTGTACAGATCTTACGCCAGGATATGACTCCAGGAGGTGGCCAGGTG
CCGCGCGTCTCAGCGGGGCGCTTGTATGCGACGCTTGGTACGCGCTCTCGCGCTGGGGCTGTCTCCCTCGAA
CCAACACCCCGCTCGGGCTCTGAGGCCAGATCCAGGTACACCTTCTGCTGCTCCCTTGGAGGCCCTAGATG
GTGGCACAGGTCCAGCCGCTGAGGGCGGGGAGTGGGTGGGCAAGATGGGAGCAGGCACCCCACTGCCCATCA
AGGCTCCCCCTCCCATCTGCTAGCACGCTCTCCAGCCCCCACTACTGTCCCTGAAGCCCCAGGGCCATGCCCCCTAC
TGCCCCCTAGAGCCATCACTATTGTCTGGAGTGGTGACGCCCTGCGAGGCGCGCTGTTGCCAGCCTGGGGCCCC
CAGGACCAACCCGGAGCCCCCTGGACCCACACAGCAGTCTTGGCCCCGGAAGATGAGGACGATGTGCTACTGG
TGCCACTGGCTGAGCCGGGGGTGGGGTAGCTGAGGCAGAGGATGAGCCACTGTCTTACCTGAGGGGACCTGGGGG
CTCTACTGAGGCTCTCCCCCTGGGGCTCTACTCATAGTGGCAACACTTTTAGAGGTGGGTGAGCTCCCCCTCC
ACCACTTCTTCCCTGTCTCTGGATTTCAGGGACTTGGTGGGCTCCCGTTGACCCCTATGTAGTGTCTATAAAGT
TAAGTGTCCCTCAGCAGGGAGAGGGCTCAGAGTCTCCTCTGTACGTGGCCATGGCCAGACACCCAGTCCCT
TCACCAACCACTGCTCCCCACGCCACCACTTTGGGTGGCTGTTTAAAGTAAAGTTCTTAGAGGATCATA
GGTCTGGACACTTCATCTTGGCAAACCTCTACCAAAGTGGCCTTAAGCACCGGAATGCCAATTAATAGAGA
CCCTCCAGCCCCCAAGGGGAGGATTTGGGCAAGAACCTGAGGTTTGGCCATCCACAATCCCTCTCAGGCGCTGG
CTCACAAAAGATGCAACAAATGCTTCTATTCCATAGCTACGGCATGTCTCAGTAAGTTGAGGTCAAATAA
GGAATCATACATCTC

09/07/2013 10:16:01

FIGURE 68

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA49631

<subunit 1 of 1, 713 aa, 1 stop

<MW: 76193, pI: 5.42, NX(S/T): 4

MLLATLLLLLLGGALAHDPRIIFPNHACEDPPAVLLEVQGTQLQRPLVRDSRTSPANCTWLIL
GSKEQTVTIRFQKLHLACGSERLTLRSPLQPLISLCEAPPSPLQLPGGNVTITYSYAGARAP
MGQGFLLSYSQDWLMCLQEEFQCLNHRCSVSAVQRCDGVDACGDGSDGSSDPFPGLTFRP
VPSLPCNVTLEDFYGVFSSPGYTHLASVSHPPQSWLLDPHDGRRRLAVRFTALDLGFGDAVH
VYDGPGPPESSRLRLSLTHFSNGKAVTVETLSGQAVVSYHTVAWSNGRGFNATYHVRGYCLP
WDRPCGLSGSLGAGEGLGERCYSEAQRCDGSWD CADGTDEEDCPGCPPGHFPCGAAGTSGAT
ACYLPADRCNYQTFCADGADERRCRHCQPGNFCRDEKCVYETWVCDGQPD CADGSEWDGCS
YVLPRKVITA AVIGSLVCGLLLVIALGCTCKLYAIRTQEYSIFAPLSRMEAEIVQQQAPPSY
GQLIAQGAIPPVEDFPTENPNDNSVLGNLRSLLQILRQDMTPGGGPGARRRQRRLMRRLVR
RLRRWGLLPRTNTPARASEARSQVTPSAAPLEALDGGTGPAREGGAVGGQDGEQAPFLPIKA
PLPSASTSPAPTTVPEAPGPIPLSLPLEPSLLSGVVQALRGRLLP SLGPPGPTRSPPGPHTAV
LALEDEDDVLLVPLAEPGVWVAEAEDEPLLT

Important features:

Signal peptide:

amino acids 1-16

Transmembrane domain:

amino acids 442-462

LDL-receptor class A (LDLRA) domain proteins

amino acids 411-431, 152-171, 331-350 and 374-393

FIGURE 69

CGAGCTGGGCGAGAAGTAGGGGAGGGCGGTGCTCCGCCGCGGTGGCGGTTGCTATCGCTTCG
CAGAACCTACTCAGGCAGCCAGCTGAGAAGAGTTGAGGGAAAGTGCTGCTGCTGGGTCTGCA
GACGCCGATGGATAACGTGCAGCCGAAAATAAAACATCGCCCCCTTCTGCTTCAGTGTGAAAGG
CCACGTGAAGATGCTGCGGCTGGCACTAACTGTGACATCTATGACCTTTTTTATCATCGCAC
AAGCCCCTGAACCATATATTGTTATCACTGGATTGAAAGTCACCGTTATCTTATTTTTTCATA
CTTTTATATGTACTCAGACTTGATCGATTAAATGAAGTGTTATTTTGGCCTTTGCTTGATAT
TATCAACTCACTGGTAACAACAGTATTCATGCTCATCGTATCTGTGTTGGCACTGATACCAG
AAACCACAACATTGACAGTTGGTGGAGGGGTGTTTGCACCTGTGACAGCAGTATGCTGTCTT
GCCGACGGGGCCCTTATTTACCGGAAGCTTCTGTTCAATCCAGCGGTCCTTACCAGAAAAA
GCCTGTGCATGAAAAAAAAAGAAGTTTGTAAATTTTATATTACTTTTTAGTTTGATACTAAGT
ATTAAACATATTTCTGTATTCTTCCAAAAAAAAAAAAAAAAAAAA

FIGURE 70

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA49645

><subunit 1 of 1, 152 aa, 1 stop

><MW: 17170, pI: 9.62, NX(S/T): 1

MDNVQPKIKHRPFCFSVKGHVKMLRLALTVTSMTFFIIAQAPEPYIVITGFEVTVILFFILL
YVLRDLRLMKWLFWPLLDIINSLVTTVFMLIVSVLALIPETTTTLTVGGGVFALVTAVCCCLAD
GALIYRKLLFNPSGPYQKKPVHEKKEVL

Important features:

Potential type II transmembrane domain:

amino acids 26-42

Other potential transmembrane domain:

amino acids 44-65, 81-101 and 109-129

Leucine zipper pattern

amino acids 78-99 and 85-106

N-myristoylation site.

amino acids 110-115

Ribonucleotide reductase large subunit protein

amino acids 116-127

FIGURE 71

GGGCGAGAAGTAGGGGAGGGCGTGTCCGCCGCGGTGGCGGTTGCTATCGTTTTGCAGAACC
TACTCAGGCAGCCAGNTGAGAAGAGTTGAGGGAAAGTGCTGCTGCTGGGTCTGCAGACGCGA
TGGATAACGTGCAGCCGAAAATAAAACATCGCCCCCTTCTGCTTCAGTGTAAGGCCACGTG
AAGATGCTGCGGCTGGCACTAACTGNGACATCTATGACCTTTTTTATNATCGCACAAAGCCCC
TGAACCATATATTGTTATCACTGGATTGAAGTCACCGTTATCTTATTTTCATACTTTTAT
ATGTACTCAGACTTGATCGATTAATGAAGTGTTATTTTGGCCTTTGCTTGATATTATCAAC
TCACTGGTAACAACAGTATTCATGCTCATCGTATCTGTGTTGGCACTGATACCAGAAACCAC
AACATTGACAGTTGGTGGAGGGGTGTTTGCACCTGTGACAGCAGTATGCTGTNTTGCCGAC

FIGURE 72

CAGCCCCGCGCGCCGGCCGAGTCGCTGAGCCGCGGCTGCCGGACGGGACGGGACCGGCTAGG
CTGGGCGCGCCCCCGGGCCCCGCCGTGGGCATGGGCGCACTGGCCCCGGGCGCTGCTGCTGC
CTCTGCTGGCCCCAGTGGCTCCTGCGCGCCGCCCGGAGCTGGCCCCCGCGCCCTTCACGCTG
CCCCTCCGGGTGGCCGCGGCCACGAACCGCGTAGTTGCGCCACCCCCGGGACCCGGGACCCC
TGCCGAGCGCCACGCCGACGGCTTGGCGCTCGCCCTGGAGCCTGCCTGGCGTCCCCCGCG
GCGCCGCCAACTTCTTGGCCATGGTAGACAACCTGCAGGGGGAAGTCTGGCCGCGGCTACTAC
CTGGAGATGCTGATCGGGACCCCCCGCAGAAGCTACAGATTCTCGTTGACACTGGAAGCAG
TAACCTTGGCGTGGCAGGAACCCCGCACTCCTACATAGACACGTACTTTGACACAGAGAGGT
CTAGCACATACCGCTCCAAGGGCTTTGACGTCACAGTGAAGTACACACAAGGAAGCTGGACG
GGCTTCGTTGGGGAAGACCTCGTCACCATCCCCAAAGGCTTCAATACTTCTTTTCTTGTCAA
CATTGCCACTATTTTTGAATCAGAGAATTTCTTTTGGCCTGGGATTAAATGGAATGGAATAC
TTGGCCTAGCTTATGCCACACTTGCCAAGCCATCAAGTTCTCTGGAGACCTTCTTCGACTCC
CTGGTGACACAAGCAAACATCCCCAACGTTTTCTCCATGCAGATGTGTGGAGCCGGCTTGCC
CGTTGCTGGATCTGGGACCAACGGAGGTAGTCTTGTCTTGGGTGGAATTGAACCAAGTTTGT
ATAAAGGAGACATCTGGTATACCCCTATTAAGGAAGAGTGGTACTACCAGATAGAAATTCTG
AAATTGGAATTGGAGGCCAAAGCCTTAATCTGGACTGCAGAGAGTATAACGCAGACAAGGC
CATCGTGGACAGTGGCACACCGCTGCTGCGCCTGCCCCAGAAGGTGTTTGATGCGGTGGTGG
AAGCTGTGGCCCGCGCATCTCTGATTCCAGAATTCTCTGATGGTTTCTGGACTGGGTCCCAG
CTGGCGTGTGGACGAATTCCGAAACACCTTGGTCTTACTTCCCTAAATCTCCATCTACCT
GAGAGACGAGAACTCCAGCAGGTCAATCCGTATCAATCCTGCCTCAGCTTTACATTACAGC
CCATGATGGGGGCCGGCCTGAATTATGAATGTTACCGATTCCGCATTTCCCATCCACAAAT
GCGCTGGTGATCGGTGCCACGGTGATGGAGGGCTTCTACGTCATCTTCGACAGAGCCCAGAA
GAGGTGGGCTTCGCAGCGAGCCCCGTGTCAGAAATTGCAGGTGCTGCAGTGTCTGAAATTT
CCGGGCCCTTCTCAACAGAGGATGTAGCCAGCAACTGTGTCCCCGCTCAGTCTTTGAGCGAG
CCCATTTTGTGGATTGTGTCTATGCGCTCATGAGCGTCTGTGGAGCCATCCTCCTTGTCTT
AATCGTCTGCTGCTGCTGCCGTTCCGGTGTGAGCGTCGCCCCCGTGACCCTGAGGTCGTCA
ATGATGAGTCCTCTCTGGTCAGACATCGCTGGAATGAATAGCCAGGCTGACCTCAAGCAA
CCATGAACCTCAGCTATTAAGAAAATCACATTTCCAGGGCAGCAGCCGGGATCGATGGTGGCG
CTTTCTCCTGTGCCCACCCGCTTCAATCTCTGTTCTGCTCCCAGATGCCTTCTAGATTACAC
TGTCTTTTGATTCTTGATTTTCAAGCTTTCAAATCCTCCCTACTTCCAAGAAAAATAATTAA
AAAAAAACTTCATTCTAA

FIGURE 73

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45493

><subunit 1 of 1, 518 aa, 1 stop

><MW: 56180, pI: 5.08, NX(S/T): 2

MGALARALLPLLAQWLLRAAPELAPAPFTLPLRVAAATNRVVAPTGPGPPTAERHADGLAL
ALEPALASPAGAAANFLAMVDNLQGDSEGRGYYLEMLIGTPPQKLQILVDTGSSNFAVAGTPHS
YIDITYFDTERSSTYRSKGFDTVTKYTQGSWTGFGEDLVITPKGFNTSFLVNIATIFESNF
FLPGIKWNGILGLAYATLAKPSSSLETFFDLSLVTQANIPNVFSMQMCGAGLPVAGSGTNGGS
LVLGGIEPSLYKGDWYTPIKEEWYYQIEILKLEIGGQSLNLDCREYNADKAIVDSGTTLLR
LPQKVFDAVVEAVARASLPEFSDGFWTGSQACWTNSETPWSYFPKISIIYLRDENSRSRFR
ITILPQLYIQPMMGAGLNYECYRFGISPSTNALVIGATVMEGFYVIFDRAQKRVGFAASPCA
EIAGAAVSEISGPFSTEDVASNCVPAQSLSEPI LWIVSYALMSVCGAILLVLLVLLLLPFRC
QRRPRDPEVVNDESSLVRHRWK

Important features:

Signal peptide:

amino acids 1-20

Transmembrane domain:

amino acids 466-494

N-glycosylation sites.

amino acids 170-173 and 366-369

Leucine zipper pattern.

amino acids 10-31 and 197-118

Eukaryotic and viral aspartyl proteases

amino acids 109-118, 252-261 and 298-310

FIGURE 74

CGCCTCCGCCCTTCGGAGGCTGACGCGCCCCGGGCGCCGTTCCAGGCCTGTGCAGGGCGGATCG
GCAGCCGCCTGGCGGCGATCCAGGGCGGTGCGGGGCTGGGCGGGAGCCGGGAGGCGCGGCC
GGCATGGAGGCGCTGCTGCTGGGCGCGGGGTGCTGTGGGCGCTTACGTGCTTGTCTACTA
CAACCTGGTGAAGGCCCCGCCGTGCGGCGGCATGGGCAACCTGCGGGGCCGACGCGCCGTGG
TCAGGGCGCCAAACAGCGGCATCGGAAAGATGACGGCGCTGGAGCTGGCGCGCCGGGGAGCG
CGCGTGGTGTGGCCTGCCGAGCCAGGAGCGCGGGGAGGCGGCTGCCTTCGACCTCCGCCA
GGAGAGTGGGAACAATGAGGTCACTTTCATGGCCTTGGACTTGGCCAGTCTGGCCTCGGTGC
GGGCCTTTGCCACTGCCCTTCTGAGCTCTGAGCCACGGTTGGACATCCTCATCCACAATGCC
GGTATCAGTTCCTGTGGCCGGACCCGTGAGGCGTTTAACTGCTGCTTCGGGTGAACCATAT
CGGTCCCTTTCTGCTGACACATCTGCTGCTGCCTTGCCTGAAGGCATGTGCCCTTAGCCGCG
TGGTGGTGGTAGCCTCAGCTGCCCACTGTGCGGGACGTCTTGACTTCAAACGCTGGACCGC
CCAGTGGTGGGCTGGCGGAGGAGCTGCGGCATATGCTGACACTAAGCTGGCTAATGTACT
GTTTGCCCCGGAGCTCGCCAAACAGCTTGAGGCCACTGGCGTCACCTGCTATGCAGCCCACC
CAGGGCCTGTGAACTCGGAGCTGTTCTCGCCATGTTCTCGGATGGCTGCGCCACTTTTG
CGCCATTGGCTTGGCTGGTGTCTCCGGGCACCAAGAGGGGTGCCAGACACCCCTGTATTG
TGCTCTACAAGAGGGCATCGAGCCCCTCAGTGGGAGATATTTTGCCAACTGCCATGTGGAAG
AGGTGCCTCCAGCTGCCCCGAGACGACCGGGCAGCCCATCGGCTATGGGAGGCCAGCAAGAGG
CTGGCAGGGCTTTGGGCTGGGGAGGATGCTGAACCCGATGAAGACCCCACTCTGAGGACTC
AGAGGCCCCATCTTCTCTAAGCACCCCCACCCCTGAGGAGCCACAGTTTCTCAACCTTACC
CCAGCCCTCAGAGCTCACCAGATTTGTCTAAGATGACGCACCGAATTCAAGCTAAAGTTGAG
CCTGAGATCCAGCTCTCCTAACCCTCAGGCCAGGATGCTTGCCATGGCACTTCATGGTCCTT
GAAAACTCGGATGTGTGTGAGGCCATGCCCTGGACACTGACGGGTTTGTGATCTTGACCTC
CGTGGTTACTTTCTGGGGCCCCAAGCTGTGCCCTGGACATCTCTTTCTCGTTGAAGGAAT
AATGGGTGATTATTTCTTCTCTGAGAGTGACAGTAACCCAGATGGAGAGATAGGGGTATGCT
AGACACTGTGCTTCTCGAAATTTGGATGTAGTATTTTCAGGCCCCACCCCTATTGATTCTG
ATCAGCTCTGAGCAGAGGCAGGGAGTTTGCAATGTGATGCACTGCCAACATTGAGAATTAG
TGAAGTATCCCTTTGCAACCGTCTAGCTAGGTAGTTAAATTACCCCATGTTAATGAAGCG
GAATTAGGCTCCCGAGCTAAGGGACTCGCCTAGGGTCTCACAGTGAGTAGGAGGAGGGCCTG
GGATCTGAACCCAAGGGTCTGAGGCCAGGGCCGACTGCCGTAAGATGGGTGCTGAGAAGTGA
GTCAGGGCAGGCAGCTGGTATCGAGGTGCCCCATGGGAGTAAGGGGACGCCTTCCGGGCGG
ATGCAGGCTGGGTCATCTGTATCTGAAGCCCTCGGAATAAAGCGCGTTGACCGCCAAA
AAAAAAAAAAAAAAAAA

FIGURE 75

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48227

<subunit 1 of 1, 377 aa, 1 stop

<MW: 40849, pI: 7.98, NX(S/T): 0

MEALLLGAGLLLGAYVLVYVNLVKAPPCGGMGNLRGRTAVVTGANSIGIGKMTALELARRGAR
VVLACRSQERGEAAAFDLRQESGNNEVIFMALDLASLASVRAFATAFLSSEPRLDILIHNAG
ISSCGRTREAFNLLLRVNHIGPFLTHLLLPCLKACAPSRVVVVASAAHCRGRLDKRLDRP
VVGWRQELRAYADTKLANVLFARELANQLEATGVTCTAAHPGPFVNSEFLRHVPGWLRPLLR
PLAWLVLRAPRGAQTPLYCALQEGIEPLSGRYFANCHVEEVPPAARDRAAHLWEASKRL
AGLGPGEDAEPEDEPQSEDSEAPSSLSTPHPEEPTVSQPYPSQSSPDLSKMTHRIQAKVEP
EIQLS

Important features:

Signal peptide:

amino acids 1-16

Glycosaminoglycan attachment site.

amino acids 46-49

Short-chain alcohol dehydrogenase family

amino acids 37-49 and 114-124

FIGURE 77

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA41404

<subunit 1 of 1, 985 aa, 1 stop

<MW: 105336, pI: 6.55, NX(S/T): 7

MGGMAQDSPPQILVHPQDQLFQGGPARMSCQASGQPPPTIRWLLNGQPLSMVPPDPHLLP
DGTLLLLLQPPARGHAHDGQALSTDGLVYTCEASNRLGTAVSRGARLSVAVLREDFQIQPRDM
VAVVGEQFTLECGPPWGHPEPTVSWWKDGKPLALQPGRHTVSGGSLLMARAESDEGTYMCV
ATNSAGHRESRAARVSIQEPQDYTEPVELLAVRIQLENTLLNPDPAEGPKPRPAVWLSWKV
SGPAAPASQSYTALFRTQTAPGGQGAPWAEELLAGWQSAELGGLHWGQDYEFKVRPSSGRARG
PDSNVLLLLRLPEKVPSAPPQEVTLKPGNGTVFVSWVPPPAENHNGIIRGYQVWSLGNTSLPP
ANWTVVGEQTQLEIATHMPGSYCVQVAAVTGAGAGEPSRPVCLLLEQAMERATQEPSEHGPW
TLEQLRATLKRPEVIATCGVALWLLLLGTAVCIHRRRRARVHLGPGLYRYTSEDAILKHRMD
HSDSQWLADTWRSTSGSRDLSSSSSLSSRLGADARDPLDCRRSLLSWDSRSRPGVPLLPDTST
FYGSLIAELPSSTPARPSPQVPAVRRLPPQLAQLSSPCSSSDSLCSRGLSSPRLSLAPAEA
WKAKKKQELQHANSSPLLRGSHSLELRACELGNRGSKNLSQSPGAVPQALVAWRALGPKLLS
SSNELVTRHLPPAPLFPHPETPTQSQQTQPPVAPQAPSSILLPAAPIPILSPCSPSPQASS
LSGPSPASSRLSSSSSLSSLGEDQDSVLTPPEEVALCLELSEGEETPRNSVSPMPRAPSPPTTY
GYISVPTASEFTDMGRTGGGVGPKGGVLLCPPRPCLTPTPSEGLANGWGSASEDNAASARA
SLVSSSDGSFLADAHFARALAVAVDSFGFGLFPREADCVFIDASSPPSPRDEIFLTPNLSLP
LWEWRPDWLEDMEVSHTQRLGRGMPPWPDPDSQISSQRSQLHCRMPKAGASPVAYS

Important features:

Transmembrane domain:

amino acids 448-467

N-glycosylation sites:

amino acids 224-227, 338-341, 367-370, 374-377, 658-661 and 926-929

N-myristoylation sites.

amino acids 47-52, 80-85, 88-93, 99-104, 105-110, 181-186, 272-277, 290-295, 355-360, 403-408, 462-467, 561-566, 652-657, 849-854 and 876-881

Phosphotyrosine interaction domain proteins

amino acids 740-753

FIGURE 78

CTCCCACGGTGTCCAGCGCCAGAAATGCGGCTTCTGGTCTGCTATGGGGTTGCCTGCTGCT
 CCAGGTTATGAAGCCCTGGAGGCCCCAGAGGAAATCAGCGGGTTCCAGAGGGACACTGTGT
 CCTGTCAGTGCACCTACAGGGAAGAGCTGAGGGACCACCGGAAGTACTGGTGCAGGAAGGGT
 GGGATCCTCTTCTCTCGCTGCTCTGGCACCATCTATGCAGAAGAAAGGCCAGGAGACAAT
 GAAGGGCAGGGTGTCCATCCGTGACAGCCGCCAGGAGCTCTCGCTATTGTGACCCCTGTGGA
 ACCTACCTTGCAGACGCTGGGGAGTACTGGTGTGGGGTGCAGAAAAACGGGGCCCCGATGAG
 TCTTTACTGATCTCTCTGTTCGTCTTCCAGGACCTGCTGTCTCCCTCCCTCTTCTCCAC
 CTTCCAGCCTCTGGCTACAAACGCTGAGCCCAAGGCAAAAGCTCAGCAAAACCAGCCCC
 CAGGATTGACTTCTCTGGGCTCTACCCGGCAGCCACCACAGCCAAGCAGGGGAAGACAGGG
 GCTGAGGGCCCCCTCATTGCCAGGGACTTCCAGTACGGGCACGAAAGGACTTCTCAGTACAC
 AGGAACCTCTCTCACCAGCGACTCTCTCTCTGAGGAGCTCCCCCCCCCATGTCAGC
 TGGACTCCACCTCAGCAGAGGACACAGTCCAGCTCTCAGCAGTGGCAGCTCTAAGCCCAGG
 GTGTCCATCCCGATGGTCCGCATACTGGCCCCAGTCTGTGTGCTGCTGAGCCTTCTGTGAGC
 CGCAGGCCGTGATCGCCTTCTGTCAGCCACCTGCTCTGTGGAGAAAAGGCTCAACAGGCCA
 CGGAGACACAGAGGAAACGAGAAGTCTGGCTCTCAGCTTGACTCGGAGGAAAAGGAAGCC
 CCTTCCAGGCCCTGAGGGGGACGTGATCTCGATGCTCCCTCCCTCCACATCTGAGGAGGA
 GCTGGGCTTCTCGAAGTTTGTCTCAGCGTAGGGCAGGAGGCCCTCTGGCCAGGCCAGCAGT
 GAAGCAGTATGGCTGGCTGGATCAGCACCGATTCGCCAAAGCTTTCCACCTCAGCCTCAGAG
 TCCAGCTGCCCGGACTCCAGGGCTCTCCCCACCCTCCCCAGGCTCTCCTCTGTCATGTTCGA
 GCCTGACCTAGAAGCGTTTGTGTCAGCCCTGGAGCCCAGAGCGGTGGCCTTGCTCTTCCGGCTG
 GAGACTGGGACATCCCTGATAGGTTCAATCCCTGGGCAGAGTACCAGGCTGCTGACCTCA
 CAGGGCCAGACAAAGGCTCAGTGGATCTGGTCTGAGTTTCAATCTGCCAGGAACTCTGGGG
 CTGATGCCCAGTGTGCGGACCTGCCTTCTCTCCACTCCAGACCCACCTTGCTCTTCCCTCCC
 TGGCGTCTCAGACTTAGTCCCACGGTCTCTGTCATCAGCTGGTGATGAAGAGGAGCATGCT
 GGGGTGAGACTGGGATTCTGGCTTCTCTTTGAACCACCTGCATCCAGCCCTTCAGGAAGCCT
 GTGAAAAACGTGATTCCTGGCCCCACCAAGACCCACCAAAACCATCTCTGGGCTTGGTGACG
 GACTCTGAATTCTAACAAATGCCAGTGACTGTGCGACTTGAGTTTGAGGGCCAGTGGGCCTG
 ATGAACGCTCACACCCCTTCAGCTTAGAGTCTGCATTTGGGCTGTGACGTCTCCACCTGCCC
 CAATAGATCTGCTCTGTCTGCGACACAGATCCACGTGGGGACTCCCTGAGGCCTGTCTAAG
 TCCAGGCCCTTGGTCAGGTGAGTGCACATTGCAGGATAAGCCAGGACCGGCACAGAAGTGG
 TTGCCCTTTNCCATTGCCCCCTGCGNCCATGCGCTTCTTGCCCTTTGGAAAAAATGATGAAGA
 AAACCTTGCGCTCTCTCTGTCTGGAAGGGTTACTTGCCATGCGGTTCTGGTGGCTAGAGA
 GAAAAGTAGAAAACCAGAGTGCACGTAGGTGTCTAACACAGAGGAGAGTAGGAACAGGGCGG
 ATACCTGAAGGTGACTCCGAGTCCAGCCCCCTGGAGAAGGGGTGGGGGTGGTGGTAAAGTA
 GCACAACTACTATTTTTTTTCTTTTTCATTATTATTGTTTTTAAGACAGAACTCTCGTGCT
 GCTGCCCAGGCTGGAGTGCAGTGGCAGCATCTGCAAACTCCGCTCTCTGGGTCAAGTGATT
 CTTTGCCTCAGCCTCCCGAGTAGCTGGGATTACAGGCACGCACCACACCTGGCTAATT
 TTTGTACTTTTAGTAGAGATGGGGTTTACCATTGTTGGCCAGGCTGGTCTTGAACCTCTGAC
 CTCAAATGAGCCTCTGCTTCAGTCTCCCAAATGCGGGATTACAGGCATGAGCCACTGTG
 TCTGGCCCTATTTCTTTAAAAAGTGAATTAAAGAGTGTTCAGTATGCAAACTTGGAAAG
 ATGGAGGAGAAAAAGAAAAGGAAAAAATGTCAACCATAGTCTCACCAGAGACTATCAT
 TATTTGGTTTGTGTAATCTCTTCCACTCTTTCTCTTCCACATAATTGGCGGTGTTCTT
 TTTACAGAGCAATTATCTGTATATACAACCTTTGTATCTGCTCTTCCACCTTATCGTTCC
 ATCACTTTATCCAGCACTTCTCTGTGTTTTACAGACCTTTTATAAATAAAATGTTTCATCA
 GCTGCATAAAAAAAAAAAAAA

03970403.10404

FIGURE 79

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA44196

<subunit 1 of 1, 332 aa, 1 stop

<MW: 36143, pI: 5.89, NX(S/T): 1

MRLLVLLWGCLLLPGYEALGPPEISGFEGDTVSLQCTYREELRDHRKYWCRKGGILFSRCS
GTIYAE EEGQETMKGRVSI RDSRQELSLIVTLWNLT LQDAGEYWCGVEKRGPD E SLLISLFV
FPGCCPPSPSPTFQPLATTRLQPKAKAQQTQPPGLTSPGLYPAATTAKQKGTGA EAPPLPG
TSQYGHERTSQTGTSPHPATSPPPAGSSRPPMQLDSTSAEDTSPALSSGSSKPRVSIPMVRI
LAPVLVLLSLLSAAGLIAFCSHLLLWRKEAQQATETQRNEKFWLSRLTAE EKEAPSQAPEGD
VISMPLHTSEEELGFSKFVSA

Important features:

Signal peptide:

amino acids 1-17

Transmembrane domain:

amino acids 248-269

N-glycosylation site.

amino acids 96-99

Fibrinogen beta and gamma chains C-terminal domain.

amino acids 104-113

Ig like V-type domain:

amino acids 13-128

FIGURE 80

TTGTGACTAAAAGCTGGCCTAGCAGGCCAGGGAGTGCAGCTGCAGGCGTGGGGGTGGCAGGA
GCCGCAGAGCCAGAGCAGACAGCCGAGAAAACAGGTGGACAGTGTGAAAGAACCAGTGGTCTC
GCTCTGTTGCCCAGGCTAGAGTGTACTGGCGTGATCATAGCTCACTGCAGCCTCAGACTCCT
GGACTTGAGAAATCCTCCTGCCTTAGCCTCCTGCATATCTGGGACTCCAGGGGTGCACCTCAA
GCCCTGTTTCTCTCTCTCTGTGTGAGTGGACCACGGAGGCTGGTGAGCTGCCTGTATCCCAA
AGCTCAGCTCTGAGCCAGAGTGGTGGTGGCTCCACCTCTGCCGCCGGCATAGAAGCCAGGAG
CAGGGCTCTCAGAAGGCGGTGGTGGCCAGCTGGGATCATGTTGTTGGCCCTGGTCTGTCTGC
TCAGCTGCCTGCTACCTCCAGTGAGGCCAAGCTCTACGGTCGTTGTGAACTGGCCAGAGTG
CTACATGACTTCGGGCTGGACGGATACCGGGGATACAGCCTGGCTGACTGGGTCTGCCTTGC
TTATTTTCAAGCGGTTTCAACGCAGCTGCTTTGGACTACGAGGCTGATGGGAGCACCAACA
ACGGGATCTTCCAGATCAACAGCCGGAGGTGGTGCAGCAACCTCACCCGAACGTCCCCAAC
GTGTGCCGGATGTACTGCTCAGATTTGTTGAATCCTAATCTCAAGGATACCGTTATCTGTGC
CATGAAGATAACCCAAGAGCCTCAGGGTCTGGGTTACTGGGAGGCCTGGAGGCATCACTGCC
AGGGAAAAGACCTCACTGAATGGGTGGATGGCTGTGACTTCTAGGATGGACGGAACCATGCA
CAGCAGGCTGGGAAATGTGGTTTGGTTCCTGACCTAGGCTTGGGAAGACAAGCCAGCGAATA
AAGGATGGTTGAACGTGAAA


```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA52187
```

<subunit 1 of 1, 146 aa, 1 stop

<MW: 16430, pI: 5.05, NX(S/T): 1

MLLALVCLLSCLLPSSSEAKLYGRCELARVLHDFGLDGYRGYSLADWVCLAYFTSGFNAAALD
YEADGSTNNGIFQINSRRWCNSLTPNVPNVCRCMYCSDLLNPNLKDTVICAMKITQEPQGLGY
WEAWRHHCHQKDLTEWVDGCDP

Important features:

Signal peptide:

amino acids 1-18

N-myristoylation site.

amino acids 67-72

Homologous region to Alpha-lactalbumin / lysozyme C proteins.

amino acids 34-58 (catalytic domain), 111-132 and 66-107

FIGURE 82

AGCCGCTGCCCCGGGCGGGCGCCGCGGCGGCACCATGAGTCCCCGCTCGTGCTGCGTTC
GCTGCGCCTCCTCGTCTTCGCGCTCTTCTCAGCCGCCGCGAGCAACTGGCTGTACCTGGCCA
AGCTGTGTCGTCGGTGGGGAGCATCTCAGAGGAGGAGACGTGCGAGAAACTCAAGGGCCTGATC
CAGAGGCAAGTGCAGATGTGCAAGCGGAACCTGGAAGTCATGGACTCGGTGCGCCGCGGTGC
CCAGCTGGCCATTGAGGAGTGCCAGTACCAGTTCCGGAACCGGCGCTGGAACCTGCTCCACAC
TCGACTCCTTGCCCGTCTTCGGCAAGGTGGTGACGCAAGGGACTCGGGAGGCGGCCCTTCGTG
TACGCCATCTCTTCGGCAGGTGTGGCCTTTGCAGTGACGCGGGCGTGCAGCAGTGGGGAGCT
GGAGAAGTGGCGCTGTGACAGGACAGTGCATGGGGTCAGCCACAGGGCTTCCAGTGGTCAG
GATGCTCTGACAACATCGCTACGCTGTGGCCTTCTCACAGTCGTTTGTGGATGTGCGGGAG
AGAAGCAAGGGGGCTCGTCCAGCAGAGCCCTCATGAACCTCCACAACAATGAGGCCGGCAG
GAAGGCCATCCTGACACACATGCGGGTGGAATGCAAGTGCCAAGGGGTGTGAGGCTCCTGTG
AGGTAAAGACGTGCTGGCGAGCCGTGCCGCCCTTCCGCCAGGTGGGTACGCACTGAAGGAG
AAGTTTGATGGTGCCACTGAGGTGGAGCCACGCCGCGTGGGCTCCTCCAGGCACTGGTACC
ACGCAACGCACAGTTCAAGCCGCACACAGATGAGGACCTGGTGTACTTGGAGCCTAGCCCCG
ACTTCTGTGAGCAGGACATGCGCAGCGCGTGTCTGGGCACGAGGGGCCGCACATGCAACAAG
ACGTCCAAGGCCATCGACGGCTGTGAGCTGCTGTGCTGTGGCCGCGGCTTCCACACGGCGCA
GGTGGAGCTGGCTGAACGCTGCAGCTGCAAATTCAGTGGTGTCTTCGTCAAGTGCCGGC
AGTGCCAGCGGCTCGTGAGTTGCACACGTGCCGATGAACCGCCTGCCCTAGCCCTGCGCCGGC
AACCACCTAGTGGCCAGGGAAGGCCGATAATTTAAACAGTCTCCACCACTACCCCAAGA
GATACTGGTTGTATTTTGTCTGTTTGGTCTTGGGTCTCATGTTATTTATTGCCGAA
ACCAGGCAGGCAACCCCAAGGGCACCAACCAGGGCCTCCCCAAAGCCTGGGCCTTTGTGGCT
GCCACTGACCAAAGGACCTTGCTCGTGCCGCTGGCTGCCCGCATGTGGCTGCCACTGACCA
CTCAGTTGTTATCTGTGTCCGTTTTTCTACTTGCAGACCTAAGGTGGAGTAACAAGGAGTAT
TACCACCACATGGCTACTGACCGTGTCTATCGGGGAAGAGGGGCCTTATGGCAGGGAATA
GGTACCGACTTGATGGAAGTCACACCCTCTGGAAAAAGAACTCTTAACTCTCCAGCACACA
TACACATGGACTCCTGGCAGCTTGAGCCTAGAAGCCATGTCTCTCAATGCCCTGAGAAAGG
GAACAAGCAGATACCAGGTCAAGGGCACAGGTTTCATTTAGCCCTTACATGGACAGCTAGA
GGTTCGATATCTGTGGGTCTTCCAGGCAAGAAGAGGGAGATGAGAGCAAGAGACGACTGAA
GTCCACCCCTAGAACCCAGCCTGCCCCAGCCTGCCCCTGGGAAGAGGAACTTAAACACTCC
CCAGACCCACCTAGGCAGGCATATAGGCTGCCATCCTGGACCAGGGATCCCGGCTGTGCCCTT
TGCAGTCATGCCGAGTCACCTTTACAGCGCTGTTCTCCATGAAACTGAAAAACACACAC
ACACACACACACACACACACACACACACACACACACACACACACAGGACACACACACACCTGCGAGA
GAGAGGGAGGAAAGGGCTGTGCCCTTGCAGTCATGCCGAGTCACCTTTACAGCACTGTTCTCT

[illegible]

<subunit 1 of 1, 351 aa, 1 stop

MSPRSCLSRLRLLVFAVFSAAASNWLYLAKLSSVSGISEEETCEKLKGLIQRQVMCKRNL
VMDSVRRGAQLAIEECQYQFRNRRWNCSTLDSLPGVFGKVTTQGTREAAFYVAISSAGVAFV
TRACSSGELEKCGCDRTVTHGVSPQGFQWSGCSNDNIAGVAFSQSFVDVRERSKGAASSRALM
NLHNNAGRAKAILTHMRVECKCHGVSGSCEVTKCWRVPPFRQVGHALKEKFDGATEVEPRR
VGSSRALVPRNAQFKPHTDEDLVYLEPSPDFCEQDMRSGVLGTRGRTCNKTSKAIDGCELLC
CGRGFHTAOVELAERCSCCKFHWCCFVKRCQCRQLVELHTCR

Signal peptide:

amino acids 1-22

amino acids 88-91 and 297-300

amino acids 206-215

amino acids 183-235, 305-350, 97-138, 53-92 and 150 -174

FIGURE 84

CGGACGCGTGGGCGGACGCGTGGGCGGACGCGTGGGCGGACGCGTGGGCTGGGTGCCTGCAT
CGCCATGGACACCACCAGGTACAGCAAGTGGGCGGCAGCTCCGAGGAGGTCCCCGGAGGGC
CCTGGGGACGCTGGGTGCACTGGAGCAGGAGACCCCTCTTCTTGGCCCTGGCTGTCTCTGGTC
ACCACAGTCTCTTGGGCTGTGATTCTGAGTATCCTATTGTCCAAGGCCTCCACGGAGCGCGC
GGCGCTGCTTGACGGCCACGACCTGCTGAGGACAAACGCCTCGAAGCAGACGGCGGCGCTGG
GTGCCCTGAAGGAGGAGGTCTGGAGACTGCCACAGCTGCTGCTCGGGGACGCAGGCGCAGCTG
CAGACCACGCGCGCGGAGCTTGGGGAGGCGCAGGCGAAGCTGATGGAGCAGGAGAGCGCCCT
GCGGGAATGCGTGAGCGCGTGACCCAGGGCTTGGCTGAAGCCGCAGGGGCGGTGAGGACG
TCCGCACTGAGCTGTTCCGGGCGCTGGAGGCCGTGAGGCTCCAGAACAACTCCTGCGAGCCG
TGCCCCACGTCGTGGCTGTCTTCCGAGGGCTCCTGCTACTTTTTCTCTGTGCCAAAGACGAC
GTGGGCGGCGCGCAGGATCACTGCGCAGATGCCAGCGCGCACCTGGTGATCGTTGGGGGCC
TGGATGAGCAGGGCTTCTCACTCGGAACACGCGTGGCCGTGGTTACTGGCTGGGCTGAGG
GCTGTGCGCCATCTGGGCAAGGTTCAAGGCTACCAGTGGGTGGACGGAGTCTCTCTCAGCTT
CAGCCACTGGAACCAGGGAGAGCCCAATGACGCTTGGGGGCGCGAGAAGTGTGTCTGATGC
TGCAACGCGGGCTGTGGAACGACGACCCGTGTGACAGCGAGAAGGACGGCTGGATCTGTGAG
AAAAGGCACAACCTGCTTGACCCCGCCAGTGCCCTGGAGCCGCGCCCATTCGAGCATGTGTA
TCCTGGGGGCTGCTCACCTCCCTGGCTCCTGGAGCTGATTGCCAAAGAGTTTTTTCTTCTCT
CATCCACCGCTGCTGAGTCTCAGAAACACTTGGCCCAACATAGCCCTGTCCAGCCAGTGCC
TGGGCTCTGGGACCTCCATGCCGACCTCATCTAACTCCACTCACGCAGACCCAACTAACC
TCCACTAGCTCCAAAATCCCTGCTCCTGCGTCCCGTGATATGCCCTCCACTTCTCTCCCTAA
CCAAGGTTAGGTGACTGAGGACTGGAGCTGTTTGGTTTTCTCGCATTTTCCACCAAACCTGGA
AGCTGTTTTTGACGCTGAGGAAGCATCAATAAATATTGAGAAATGAAAAA

09976403-101501

FIGURE 85

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56352

<subunit 1 of 1, 293 aa, 1 stop

<MW: 32562, pI: 6.53, NX(S/T): 2

MDTTRYSKWGGSSSEEVPGGPWGRVWHWSRRPLFLALAVLVTTLWAVILSILLSKASTERAA
LLDGHDLLRTNASKQTAALGALKEEVGDCHSCCSGTQAQLQTTTAEELGEAQAKLMEQESALR
ELRERVTOGLAEAGRGREDVRTELFRALEAVRLQNNSCEPCPTSWLSFEGSCYFFSVPKTTW
AAAQDHCADASAHLVIVGGLDEQGFLTRNTRGRGYWLGLRAVRHLGKVQGYQWVDGVSLSFS
HWNQGEFPNDAWGRENCVMMLHTGLWNDAPCDSEKDGWICEKRHNC

Important features:

Type II transmembrane domain:

amino acids 31-54

N-glycosylation sites.

amino acids 73-76 and 159-162

Leucine zipper pattern.

amino acids 102-123

N-myristoylation sites.

amino acids 18-23, 133-138 and 242-247

C-type lectin domain signature.

amino acids 264-287

FIGURE 86

GCCAGGGGAAGAGGGTGATCCGACCCGGGGAAGGTCGCTGGGCAGGGCGAGTTGGGAAAGCG
GCAGCCCCCGCCGCCCGCAGCCCTTCTCCTCCTTTCTCCACGTCCTATCTGCCTCTCG
CTGGAGGCCAGGCCGTGCAGCATCGAAGACAGGAGGAACTGGAGCCTCATTTGGCCGGCCCGG
GGCGCCGGCCTCGGGCTTAAATAGGAGCTCCGGGCTCTGGCTGGGACCCGACCGCTGCCGGC
CGCGCTCCCGCTGCTCTCTGCCGGGTGATGGAAGAAACCCAGCCCGGCCGCCCTGGGCAAG
GCCCTCTGCGCTCTCCTCTGGCCACTCTCGGCGCCGCCGGCCAGCCTCTTGGGGGAGAGTC
CATCTGTTCCGCCAGAGCCCCGGCCAAATACAGCATCACCTTCACGGGCAAGTGGAGCCAGA
CGGCCTTCCCCAAGCAGTACCCCTGTTCCGCCCCCTGCGCAGTGGTCTTCGCTGCTGGGG
GCCGCGCATAGCTCCGACTACAGCATGTGGAGGAAGAACCAGTACGTACGTAAACGGGCTGCG
CGACTTTGCGAGCGCGGCGAGGCTGGGCGCTGATGAAGGAGATCGAGGCGGCGGGGAGG
CGTGCAGAGCGTGACAGAGGTGTTTTCGGCGCCCGCGTCCCCAGCGGCACCGGGCAGACG
TCGGCGGAGCTGGAGGTGCAGCGCAGGCACTCGCTGGTCTCGTTTGGTGCATCGTGCC
CAGCCCCGACTGGTTCTGGGCGTGGACAGCCTGGACCTGTGCGACGGGGACCGTTGGCGGG
AACAGGCGGCGCTGGACCTGTACCCCTACGACGCCGGGACGGACAGCGGCTTCACCTTCTCC
TCCCCCAACTTCGCCACCATCCCGCAGGACACGGTGACCGAGATAACGTCTCCTCTCCAG
CCACCCGGCCAACTCCTTCTACTACCCGCGGCTGAAGGCCCTGCCTCCCATCGCCAGGGTGA
CACTGCTGCGGCTGCGACAGAGCCCCAGGGCCTTCATCCCTCCCGCCCCAGTCTTCCCAGC
AGGGACAATGAGATTGTAGACAGCGCCTCAGTTCAGAAACGCCGCTGGACTGCGAGGTCTC
CCTGTGGTCTGCTCCTGGGGACTGTGCGGAGGCCACTGTGGGAGGCTCGGGACCAAGAGCAGGA
CTCGCTACGTCCGGGTCCAGCCCGCCAACAACGGGAGCCCTGCCCCGAGCTCGAAGAAGAG
GCTGAGTGCGTCCCTGATAAAGTTCGTCTAAGACCAGAGCCCCGAGCCCCGGGGCCCCCG
GAGCCATGGGGTGTCGGGGGCTCCTGTGACGGCTCATGCTGCAGGCGGCCGAGGGCACAGGG
GGTTTCGCGCTGCTCCTGACCGCGGTGAGGCCGCGCCGACCATCTCTGCACTGAAGGGCCCT
CTGGTGGCCGGCACGGGCATTTGGGAAACAGCCTCCTCCTTTCCCAACCTTGCTTCTTAGGGG
CCCCCGTGTCCCGTCTGCTCTCAGCCTCCTCCTCTGCAGGATAAAGTCATCCCCAAGGCTC
CAGCTACTCTAAATTATGTCTCCTTATAAGTTATTGCTGTCCAGGAGATTGCTCTTATCG
TCCAGGGGCTGGCTCCACGTGGTTGCAGATACCTCAGACCTGGTGTCTTAGGCTGTGCTG
AGCCCACTCTCCGAGGGCGCATCAAGCGGGGGCCACTTGAGAAGTGAATAAATGGGGCGG
TTTCGGAAGCGTCAGTGTTCATGTTATGGATCTCTCTGCGTTTGAATAAAGACTATCTCT
GTTGCTCACAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

Figure 1 *Diagram illustrating the experimental design. The study was divided into two main phases: Pre-Test and Main Test. The Pre-Test phase involved a familiarization session (Familiarization) and a pre-test session (Pre-Test). The Main Test phase involved a series of trials (Trials) and a post-test session (Post-Test). The diagram shows the sequence of events and the timing of the sessions.*

```
><subunit 1 of 1, 331 aa, 1 stop
```

MENPSPAALGKALCALLLATLGAAGQPLGGESIC SARAPAKYSITFTGKWSQTAFPKQYPL
 FRPPAQWSSLLGAHSSDYSMWRKNQYVSNGLRDFAEERGEAWALMKEIEAAGEALQSVHEVF
 SAPAVPSGTGQTSAELEVQRRHSLVSFVVRIVPSPDWFGVDSLDLCDGDRWREQAALDLYP
 YDAGTDSGFTFSSPNFATIPQDTVTEITSSSPSHPANSFYYPRLKALPPIARVTLRLRQSP
 RAFIPPAPVLPSPRDNEIVDSASVPETPLDCEVSLWSSWGLCGHGCGRLGTKSRTRYVRVQPA
 NNGSPCPELEEEEAECVPDNCV

Signal peptide:

amino acids 1-26

FIGURE 88

GGCGGCGTCCGTGAGGGGCTCCTTTGGGCAGGGGTAGTGTGTTGGTGTCCCTGTCTTGCGTGA
TATTGACAAACTGAAGCTTTCTGCACCACTGGACTTAAGGAAGAGTGTACTCGTAGGCGGA
CAGCTTTAGTGGCCGGCCGGCCGCTCTCATCCCCGTAAGGAGCAGAGTCCCTTTGTACTGAC
CAAGATGAGCAACATCTACATCCAGGAGCCTCCACGAATGGGAAGGTTTTATTGAAAACTA
CAGCTGGAGATATTGACATAGAGTTGTGGTCCAAAGAAGCTCCTAAAGCTTGCAAAATTTT
ATCCAACTTTGTTTTGAAGCTTATTATGACAATACCATTTTTCATAGAGTTGTGCCTGGTTT
CATAGTCCAAGGCGGAGATCCTACTGGCACAGGGAGTGGTGGAGAGTCTATCTATGGAGCGC
CATTCAAAGATGAATTTCAATCACGGTTGCGTTTTTAATCGGAGAGGACTGGTTGCCATGGCA
AATGCTGGTTCTCATGATAATGGCAGCCAGTTTTTCTTCACACTGGGTCGAGCAGATGAACT
TAACAATAAGCATACCATCTTTGGAAAGGTTACAGGGGATACAGTATATAACATGTTGCGAC
TGTCAGAAAGTAGACATTGATGATGACGAAAGACCACATAATCCACACAAAATAAAAAAGCTGT
GAGGTTTTGTTTAACTCCTTTTGATGACATCATTCCAAGGGAAATTAAAAGGCTGAAAAAAGA
GAAACCAGAGGAGGAAGTAAAGAAATTGAAACCCAAAGGCACAAAAAATTTTAGTTTACTTT
CATTTGGAGAGGAAGCTGAGGAAGAAGAGGAGGAAGTAAATCGAGTTAGTCAGAGCATGAAG
GGCAAAAGCAAAGTAGTCATGACTTGCCTTAAGGATGATCCACATCTCAGTTCTGTTCCAGT
TGTAGAAAGTGAAAAAGGTGATGCACCAGATTTAGTTGATGATGGAGAAGATGAAAGTGCAG
AGCATGATGAATATATTGATGGTGTGATAAAGAACCTGATGAGAGAAAGAATTGCCAAAAAA
TTAAAAAAGGACACAAGTGCGAATGTTAAATCAGCTGGAGAAGGAGAAGTGGAGAAGAAATC
AGTCAGCCGAGTGAAGAGCTCAGAAAAGAAGCAAGACAATTAAAACGGGAAGCTCTTAGCAG
CAAAACAAAAAAGTAGAAAAATGCAGCAAAACAAGCAGAAAAAAGAAGTGAAGAGGAAGAA
GCCCTCCAGATGGTGCTGTTGCCGAATACAGAAGAGAAAAGCAAAAGTATGAAGCTTTGAG
GAAGCAACAGTCAAAGAAGGGAACTTCCCGGAAGATCAGACCCTTGCACTGCTGAACCAGT
TTAAATCTAAACTCACTCAAGCAATTGCTGAAACACCTGAAAATGACATTCTTGAAACAGAA
GTAGAAGATGATGAAGGATGGATGTACATGTAATTCAGTTTGAGGATAAAAGCAGAAAGT
GAAAGATGCAAGCATGCAAGACTCAGATACATTTGAAATCTATGATCCTCGGAATCCAGTGA
ATAAAAGAAGGAGGGAAGAAAGCAAAAGCTGATGAGAGAGAAAAAGAAAGAAGATTAAAAAT
GAGAATAATGATAACCAGAACTTGCTGGAATGTGCCTACAATGGCCTTGTAACAGCCATTG
TTCCCAACAGCATCACTTAGGGGTGTGAAAAGAAGTATTTTTGAACCTGTTGTCTGGTTTTG
AAAAACAATTATCTTGTGTTTTCGAAATGTGGAATGATGTAAGCAAATGCTTTTGGTTACTGG
TACATGTGTTTTTCTAGCTGACCTTTTATATTGCTAAATCTGAAATAAAATAAATTTCTCT
TCCACAAAAA

FIGURE 89

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50919
><subunit 1 of 1, 472 aa, 1 stop
><MW: 53847, pI: 5.75, NX(S/T): 2
MSNIYIQEPPTNGKVLTKTAGDIDIELWSKEAPKACRNFIQLCLEAYDNTIFHRVVPGFI
VQGGDPTGTGSGGESIYGAPFKDEFHSRLRFNRRGLVAMANAGSHDNGSQFFFTLGRADELN
NKHTIFGKVTGDTVYNMLRLSEVDIDDDERPHNPHKIKSCEVLFPFDDIIPREIKRLKKEK
PEEEVKKLKPKGTKNFSLLSFGEEABEEEEEEVNRSQSMKGKSKSSHLLKDDPHLSSVPV
ESEKGDAPDLVDDGEDESAEHDEYIDGDEKNLMRERIAKCLKKDTSANVKSAGEGEVEKKSV
SRSEELRKEARQLKRELLAAKQKKVENAAKQAEKRSEEEAAPDGAVA EYRREKQKYEALRK
QQSKKGT SREDQTLALLNQFKSLTQAIAETPENDIPETEVEDDEGWMSHVLQFEDKSRKVK
DASMQSDSTFEIYDPRNPVNKRREESKKLMREKKERR
```

Important features:

Signal peptide:

amino acids 1-21

N-glycosylation sites.

amino acids 109-112 and 201-204

Cyclophilin-type peptidyl-prolyl cis-trans isomerase signature.

amino acids 49-66

Homologous region to Cyclophilin-type peptidyl-prolyl cis-trans isomerase

amino acids 96-140, 49-89 and 22-51

FIGURE 90

CGCCGCCGCTTGGGGCTGGAAGTTCCCGCCAGGTCCGTGCCGGGCGAGAGAGATGCTGCCCGG
 CCCGCTCGGCTTTGAGGCGAGAGAAGTGTCCCAGACCCATTTTCGCTTGTGACGGCGTCG
 AGCCCTGGCCAGACATGTCACACAGGGTTCTCCTTCGGGTCCGGGACTCTGGGCTCCACCACC
 GTGGCCGCGCGGGGACCAGCACAGGCGGCGTTTCTCCTTCGGAACGGGAACGCTCTAGCAA
 CCCTTCTGTGGGGCTCAATTTTGGAAATCTTGAAGTACTTCAACTCCAGCAACTACATCTG
 CTCCTTCAAGTGGTTTTTGGAAACCGGGCTCTTTGGATCTAAACCTGCCACTGGGTTCACTCTA
 GGAGGAACAAATA CAGGTGCCCTTGCACACCAAGAGGCCTCAAGTGGTCAACAAATATGGAAC
 CCTGCAAGGAAAAAGATGCATGTGGGGAAGACACCCATCCAAGTCTTTTAGGAGTCCCTT
 TCTCCAGACCTCCTCTAGGTATCCTCAGGTTTGCACCTCCAGAACCCCCGGAGCCCTGGAAA
 GGAATCAGAGATGCTACCACCTACCCGCTGGATGGAGTCTCGTCTGTGCCAGGCTGGAG
 TGCAGTGGCACGATCTCGGCTCACTGCAACCTCCGCTCCCGGGTTCAAGCGAGTCTCCTGC
 CTCAGCTCTGAGTGTCTGGGGCTACAGGTGCCTGCAGGAGTCTGGGGCCAGCTGGCCTCG
 ATGTACGTGACGACGCGGGAACGGTACAAGTGGCTGCGCTTCAGCGAGGACTGTCTGTACCT
 GAACGTGTACGCGCCGGCGCGCGCGCCGGGATCCCCAGCTGCCAGTGATGGTCTGGTTCC
 CGGGAGGCGCCTTCATCGTGGGCGCTGCTTCTTCGTACGAGGGCTCTGACTTGGCGCGCCGCG
 GAGAAAGTGGTGTGGTGTCTTCTGCAGCACAGGCTCGGCTCATCTTCGGCTTCTGAGCACGGA
 CGACAGCCACGCGCGCGGGAACCTGGGGGCTGTCTGGACAGATGGCGGCTCTGCGCTGGGTGC
 AGGAGAACATCGCAGCCTTCGGGGGAGACCCAGGAAATGTGACCTGTTTCGGCCAGTCCGCG
 GGGGCCATGAGCATCTCAGGACTGATGATGTACCCCTAGCCTCGGGTCTCTTCCATCGGGC
 CATTTCCAGAGTGGCACCGGCTTATTCAGACTTTTCATCACTAGTAACCCACTGAAAGTGG
 CCAAGAAAGTTGCCACCTGGCTGGATGCAACCACAACAGCACACAGATCCTGGTAAACTGC
 CTGAGGGCACTATCAGGGACCAAGGTGATGCGTGTGTCCAAAGATGAGATTCTCCAAC
 GAACTTCCAGAGAGACCCGGAAGAGATTATCTGGTCCATGAGCCCTGTGGTGGATGGTGTGG
 TGATCCAGATGACCTTTGGTGTCTCCTGACCCAGGGGAAGGTTTTCATCTGTGCCCTACCTT
 CTAGGTGTCAACAACCTGGAATTCAATTGGCTCTTGCTTATAATATCACCAGGAGCAGGT
 ACCACTTGTGGTGGAGGAGTACCTGGACAATGTCAATGAGCATGACTGGAAGATGCTACGAA
 ACCGTATGATGGACATAGTTCAAGATGCCACTTTCGTGTATGCCACACTGCAGACTGCTCAC
 TACCACCGAGAAACCCCAATGATGGGAATCTGCCCTGTGGCCACGCTACAACAAGGATGAA
 AAGTACCTGCAGCTGGATTTTACCAACAAGAGTGGGCATGAAGCTCAAGGAGAAGAAGATGGC
 TTTTGGATGAGTCTGTACAGTCTCAAAGACCTGAGAAGCAGAGGCAATTCTAAGGGTGGC
 TATGCAGGAAGAGCCAAAGAGGGGTTTGGCCCCACCATCCAGGCCCTGGGGAGACTAGCCA
 TGGACATACCTGGGGACAAGAGTTCTACCAACCCAGTTTAGAACTGCAGGAGCTCCCTGCT
 GCCTCCAGGCCAAAGCTAGAGCTTTTGCTGTGTGTGGGACCTGCACTGCCCTTTCCAGCC
 TGACATCCCATGATGCCCTCTACTTCACTGTTGACATCCAGTTAGGCCAGGCCCTGTCAAC
 ACCACACTGTGCTCAGCTCTCCAGCCTCAGGACAACCTCTTTTTCCTTCTTCAAATCCT
 CCCACCTTCAATGTCTCCTTGTGACTCCTTCTTATGGAGGTCGACCCAGACTGCCACTGC
 CCCTGTCACTGCAACCCAGCTTGGCATTTACCATCCATCCTGCTCAACCTTGTTCCTGTCTGT
 TCACATTGGCCCTGGAGGCCTAGGGCAGGTTGTGACATGGAGCAAACTTTTGGTAGTTTGGGA
 TCTTCTCTCCACCCACACTTATCTCCCCCAGGGCCACTCAAAGTCTATACACAGGGGTG
 TCTCTTCAATAAAGAAGTGTGATTAGAAAAA

CGCGGCTTGGGGCTGGAAGTTCCCGCCAGGTCCGTGCCGGGCGAGAGAGATGCTGCCCGG

FIGURE 91

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA44179

<subunit 1 of 1, 545 aa, 1 stop

<MW: 58934, pI: 9.45, NX(S/T): 4

MSTGFSFGSGTLGSTTVAAGGTSTGGVFSFGTGTSSNPSVGLNFGNLGSTSTPATTSA PSSG
FGTGLFSGK PATGFTLG GTNTGALHTKRPQVVTKYGT LQ GKQMHVGKTP IQVFLGVPFSRPP
LGILRFAPPEPPEPWKGIRDAT TYPPGWQLALSPGWSAVARSRLTATSASRVQASLLPQPLS
VWGYRCLQESWGQLASMYVSTRERYKWLRFSEDCLYLNVYAPARAPGDPQLPVMVWFPGGAF
IVGAASSYEGSDLAAREKVVLVFLQHRLGIFGFLSTDDSHARGNWGLLDQMAALRWVQENIA
AFGGDPGNVT LFGQSAGAMSI SGLMMSPLASGLFHRAISQSGTALFRLFITSNPLKVAKKVA
HLAGCNHNSTQILVNCLRALSGTKVMRVS NKMRFLQLNFORDP EEI IWSMSPVVDGVVIPDD
PLVLLTQGKVSSVPYLLGVNNLEFNWLLPYNITKEQVPLVVEEYLDNVNEHDWKMLNRNMMMD
IVQDATFVYATLQTAHYHRET PMMGICPAGHATTRMKSTCSWILPQEWA

Important features:

Signal peptide:

amino acids 1-29

Carboxylesterases type-B serine active site.

amino acids 312-327

Carboxylesterases type-B signature 2.

amino acids 218-228

N-glycosylation sites.

amino acids 318-321, 380-383 and 465-468

09978403 101601
105101-2048260

GAGAACAGGCGCTGTCTCAGGCAGGCGCTGCGCCTCCTATGCGGAGATGTGCTACTGCCACTGCT
GCTGTCTCTCGTCTGCTGGGCGGCTCCAGGCTATGGATGGGAGATCTTGGATACAGGTGCGAGG
AGTCAGTGAATGGTGCGCGGAGGGCTGTGCATCTCTGTGCCCTGCTCTTTCTCTCCATCCCCGA
CAAGACTGGACAGGGTCTACCCAGCTTATGGCTACTGGTTCAAAGCAGTGACTGAGACAAC
CAAGGGTGCTCCTGTGGCCACAACACAGAGTCGAGAGGTGGAATGAGCAGCCGGGGCC
GATTCAGCTCACTGGGGATCCCGCAGGGGAATGCTCCTTGGTGATCAGAGACGGCGAG
ATCGAGACTGAGTCACAGTACTTCTTTCCGGTGAGAGGAAGCACTATGTGACATATAATT
CATGAACGATGGGTTCTTTCTAAAAGTAACAGTGCTCAGCTTCACGCCCAGACCCAGGACC
ACAACACCGCACTCACCCTGCCATGTGGACTTCTCCAGAAGGGTGTGAGCGCACAGAGGACC
GTCCGACTCCGTGTGGCTATGCCCCAGAGACTTGTATTGACATTTCAGTGAACAACAC
GCCAGCCTGGAGCCCCAGCCCCAGGGAATGTCCATACCTGGAAGCCCAAAAGGCCAGT
TCCTGCGGCTCCTCTGTGCTGTGACAGCCAGCCCCCTGCCACACTGAGCTGGGTCTGCGAG
ACAAGAGTCTCTCTCTGCTCCATCTCCCTGGGGCCCTAGACCCCTGGGGCTGGAGCTGCCCGG
GGTGAAGCTGGGGATCTCAGGGCTCTACCTGCGCAGCGGAGAACAGGCTTGGCTTCCGAGC
AGCGAGCCCTGGACCTCTCTGTGCAGTATCTCCAGAGAACCTGAGAGTGATGGTTTCCCAA
GCAAAACAGGACAGTCTCGGAAAACCTTGGGAAACCGCACTGCTCTCCAGTATGGAGGGCCA
AAGGCTGTGCTGGTCTGTGTACACACAGCAGCCCCCAGCGAGCTGAGCTGGAGCCAGA
GGGGACAGGTTCTTGAGCCCTCCAGCCCTCAGACCCCGGGGTCTGGAGCTGCCTCGGGTT
CAAGTGAGACGAAGGAGAGTTCACCTGCCACGCTCGCGACCCACTGGGCTCCCAGCAGCT
CTCTCTCAGCCTCTCCCGTGCACTATAAGAAAGGACTCATCTCAACGGCACTTCTCCACGGAG
CGTTTCTGGGAAATCGGCATCAGCGCTCTCTTTCTCTGCTGGCCCTGATCATCATGAAG
ATTCTACCGAAGAGACGGACTCAGACAGAAACCCGAGGCCAGGTTCTCCCGGCACAGCAC
GATCCTGGATTACATCAATGTGTGCCGACGCTGGCCCCCTGGCTCAGAAGCGGAATCAGA
AAGCCACACAAAACAGTCTCGGACCCTCTCCACAGGTCGCTCCTCCAGCAATCAAAG
AAGAACACAGAAAAGCAGTATCAGTTGCCAGGTTTCCAGAAACCAATCATCCACTCAAGC
CCCAGAAATCCCAGGAGAGCCAAAGAGGAGTCCATTATGCCACGCTCAACTTCCCAGGCGTCA
GACCCAGGCTGAGGCCCGGATGCCCAAGGCGACCCAGCGGATTTATGAGAAGTCAAGTTCA
CAATGAAGGGTCTCTTAGGCTTTAGGACTGGGACTTCGGCTAGGGAAGGTAAGAGTAAGAG
GTTGAAGATAACAGAGTGCAAAGTTTCTTCTCTCCTCTCTCTCTCTCTCTCTCTCTCTCTCT
CTCTCTTCTCTCTCTCTTTTAAAAAACATCTGGCCAGGGCAGCAGTGGGCTCAGCGCTGTAATC
CCAGCACTTTGGGAGGTTGAGGTGGGCAGATCGCTGAGTTCGGGAGTTCGAGACCGCCCTG
GCCAACTTGGTGAAACCCCGTCTCTACTAAAAATACAAAAATTAGCTGGGCATGGTGGCAGG
CGCTGTGAATCTTACTACTTGGGAGCTGAGGCAGGAGAACTACTGAACTGGGAGACGG
AGGTTGCAGTGAGCCGAAGTACACCACTTGCCAGCCAGGCTGGGCAACAAAGCGAGACTCCA
TCTCAAAAAAAAATCTCCAAATGGGTTGGGTGTCTGTAATCCAGCACTTTGGGAGGCTA
AGGTGGGTGGATTGCTTGAAGCCAGGAGTTGCAGACAGCGCTGGGCAACATGGTGAAACCCC
ATCTCTACAAAAAATACAAAACATAGCTGGGCTTGGTGGTGTGCTGTGATGCTCCAGCTGT
CAGACATTAAACAGAGCAACTCATCTGGAATAGGAGCTGTAATAAATAGGCTGAGACC
TACTGGGCTGCATTCTCAGACAGTGGAGGCATTCTAAGTCACAGGATGAGACAGGAGGTCG
TACAAGATACAGGTCATAAAGACTTTGTGCTATAAAGACAGTTGCGATAAAGAGCACAACCA
ATCCCACAAAACCAAGTTGGCCACGAGAGTGACTCTGTGCTCTCACTGCTCACTCACTCT
GCAGCACCATGACAGCTTTACAAATGCCATGGCAACATCAGGAAGTTACCCGATATGTCCCA
AAAGGGGAGGAATGAATAATCCACCCCTTTGTAGCAAAATAGCAAGAAATAACCATAAAA
GTGGGCAACCGAGCAGCTCTAGGCGCTGCTCTGTGCTATGGAGTAGCCATTCTTTTGTCTCTT
TACTTTCTTAATAAATCTGCTTTTCACTTTAAAAAAA

FIGURE 93

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA54002

><subunit 1 of 1, 544 aa, 1 stop

><MW: 60268, pI: 9.53, NX(S/T): 3

MLLPLLLSSLLGGSQAMDGRFWIRVQESVMVPEGLCISVPCSFYPRQDWTGSTPAYGYWFK
AVTETTTKAPVATNHQSREVEMSTRGRFQLTGDEPAKGNCSLVIRDAQMQDESQYFFRVERGS
YVTYNFMNDGFFLKVTVLSFTPRPQDHNTDLTCHVDFSRKGVSAQRTVRLRVAYAPRDLVIS
ISRDNTPALEPQPQGNVPYLEAQKGQFLRLLCAADSQPPATLSWVLQNRVLSSSHPWGPRPL
GLELPGVKAGDSGRYTCRAENRLGSQQRALDLSVQYPPENLRVMVSQANRTVLENLGNLSTL
PVLEGQSLCLVCVTHSSPPARLSWTQRGQVLSPSQPSDPGVLELPRVQVEHEGEFTCHARHP
LGSQHVSLSLSVHYKKGLISTAFSNGAFLGIGITALLLFLCLALIIMKILPKRRTQTETPRPR
FSRHSTILDYINVVPTAGPLAQKRNQKATPNSPRTPPPPGAPSPESKKNQKKQYQLPSFPEP
KSSTQAPESQESQEELHYATLNFPGVRPRPEARMPKGTQADYAEVKFQ

Important features:

Signal peptide:

amino acids 1-15

Transmembrane domain:

amino acids 399-418

N-glycosylation site.

amino acids 100-103, 297-300 and 306-309

**Immunoglobulins and major histocompatibility complex proteins
signature.**

amino acids 365-371

FIGURE 94

TGAAGAGTAATAGTTGGAATCAAAAGAGTCAACGCAATGAACGTGTTATTTACTGCTGCGTTT
TATGTTGGGAATTCTCTCTATGGCCTTGTCCTGGAGCAACAGAAAACCTCTCAAACAAAGA
AAGTCAAGCAGCCAGTGCATCTCATTGAGAGTGAAGCGTGGCTGGGTGTGGAACCAATTT
TTTGTACCAGAGGAAATGAATACGACTAGTCATCACATCGGCCAGCTAAGATCTGATTTTGA
CAATGGAAACAATCTTTCCAGTACAAGCTTTTGGGAGCTGGAGCTGGAACTACTTTTATCA
TTGATGAAAGAACAGGTTGACATATATGCCATACAGAAGCTTGATAGAGAGGAGCGATCCCTC
TACATCTTAAGAGCCAGGTAAATAGACATCGCTACTGGAAGGGCTGTGGAACTCGTAGCTTGA
GTTTGTCTATCAAAGTTTCGGATATCAATGACAATGAACCAAAATTCCTAGATGAACCTTATG
AGGCCATTGTACCAAGAGATGTCTCCAGAAGGAACATTAGTTATCCAGGTGACAGCAAGTGAT
GCTGACGATCCCTCAAGTGGTAATAATGCTCGTCTCCTCTACAGCTTACTTCAAGGCCAGCC
ATATTTTCTGTTGAACCAACAACAGGAGTCTATAAGAATATCTTCTAAAATGGATAGAGAAC
TGCAAGATGAGTATTGGGTAACTCATTCAAGCCAAGGACATGATTGGTCAGCCAGGAGCGTTG
TCTGGAACAACAAGTGTATTAAATTAACCTTTCAGATGTTAATGACAATAAGCCTATATTTAA
AGAAAGTTTATACCGCTTGACTGTCTCTGAATCTGCACCCACTGGGACTCTCTATAGGAACAA
TCATGGCATATGATAAATGACATAGGAGAGAATGCAGAAATGGATTACAGCATTTGAAGAGGT
GATTTCGAAACATTTGACATTATTACTAATCATGAACTCAAGAAGGAATAGTTATATTTAA
AAAGAAAGTGGATTTTGAGCACCAGAACCACTACGGTATTAGAGCAAAAGTTAAAAACCATC
ATGTTCTGAGCAGCTCATGAAGTACCACACTGAGGCTTCCACCATTTCATTAAGATCCAG
GTGGAAGATGTTGATGAGCCTCTCTTTTCTCTCTTCCATATTATGTAATGAGTTTGTGA
AGAAACCCACAGGGATCATTTGTAGGCGTGGTGTCTGCCACAGACCCAGACAATAGGAAAT
CTCCTATCAGGTATTCTATTACTAGGAGCAAAAGTGTCAATATCAATGATAATGGTACAATC
ACTACAAGTAACCTCACTGGATCGTGAAATCAGTCTGGTACAACCTTAAGTATTACAGCCAC
AGAAAAATACAATATAGAACAGATCTCTTCGATCCCACTGTATGTGCAAGTTCTTAAACATCA
ATGATCATGCTCCTGAGTTCTCTCAATATATGAGACTTATGTTTGTGAAATGCAGGCTCT
GGTCAGGTAATTCAGACTATCAGTGCAGTGGATAGAGATGAATCCATAGAAGAGCCACATT
TTACTTTAATCTATCTGTAGAAGACCTAACAATTCAGTTTTACAATCATAGATAATCAAG
ATAACACAGCTGTCATTTGACTAATAGAACTGGTTTTAACCTTCAAGAAGAACCTGTCTTC
TACATCTCCATCTTAATTGCCGACAATGGAATCCCGTCACTTACAAGTACAACACCCCTTAC
CATCCATGTCTGTGACTGTGGTGACAGTGGGAGCACACAGACCTGCCAGTACCAGGAGCTTG
TGTCTTCCATGGGATTCAAGACAGAAGTTATCATTGCTATTCTCATTGTCATTATGATCATA
TTTGGGTTTATTTTTTGTACTTTGGGTTTAAAAACAACGGAGAAAAACAGATTCTATTTCTGA
GAAAAGTGAAGATTTTCAGAGAGAAATATTTCCAATATGATGATGAAGGGGGTGGAGAAGAG
ATACAGAGGCCTTTGATATAGCAGAGCTGAGGAGTAGTACCATAATGCGGGAAACGCAAGACT
CGGAAAACCAAGCGCTGAGATCAGGAGCCTATACAGGCAGTCTTTGCAAGTTGGCCCCGA
CAGTGCCATATTAGGAAATTCATTCTGGAAAAGCTCGAAGAAGCTTAATACTGATCCGTGTG
CCCCTCCTTTTGATTCCCTCCAGACCTACGCTTTTGAGGGAAACAGGGTCATTAGCTGGATCC
CTGAGCTCCTTAGAATCAGCAGTCTCTGATCAGGATGAAAGCTATGATTACCTTAATGAGTT
GGGACCTCGCTTTAAAGATTAGCATGCATGTTTGGTCTGCGAGTGCAGTCAATAATATTAGG
GCTTTTTACCATCAAAATTTTAAAAGTGCTAATGTGTATTTCGAACCCAATGGTAGTCTTAA
AGAGTTTTGTGCCCTGGCTCTATGGCGGGGAAAGCCCTAGTCTATGGAGTTTCTGATTTCC
CTGGAGTAAATACTCCATGGTTATTTTAAAGCTACCTACATGCTGTCTATTGAACAGAGATGTG
GGGAGAAATGTAAACAATCAGCTCAGGATCAATAACAGCAATTTGAAGTAAATAATG
TAGGAAGATATTTAAAGTAGATGAGAGGACACAAGATGTAGTCTGATCCTTATGCGATTATAT
CATTTATTTACTTAGGAAAGAGTAAAAATACAAACGAGAAAAATTTAAAGGAGCAAAATTTG
CAAGTCAAAATGAAATGTACAAATCGAGATAACATTACATTTCTATCATTTGACATGGAAA
ATTGAAAATGTATAGTCAGAGAAATTTTCATGAATTAATCCATGAAGTATTGTTTCTTTAT
TTAAA

00072403 101501

FIGURE 95

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA53906
><subunit 1 of 1, 772 aa, 1 stop
><MW: 87002, pI: 4.64, NX(S/T): 8
MNCYLLLRFLMGIPLLLWPCLGATENSQTKKVKQPVRSHLRVVRGWVWNQFFVPEEMNTTSHH
IGQLRSDLDNGNNSFYQYKLLGAGAGSTFIIDERTGDIYAIQKLDREERSLYILRAQVIDIAT
GRAVEPESEFVIKVSNDINDNEPKFLDEPYEAIVPEMSPEGTLVIQVTASDADDPSSGNNARL
LYSLQGQPYFSVEPTTGVIRISSKMDRELQDEYWVIIQAKDMIGQPGALSGTTSVLIKLS
VNDNKPIFKESLYRLTVSESAPTGTSGITIMAYDNDIGENAEMDYSIEEDDSQTFDIIITNHE
TQEGIVILKKKVDFEHQNHYGIRAKVKNHHVPEQLMKYHTEASTTFIKIQVEDVDEPPLFLL
PYYVFEVFEETPQGSFVGVSATDPDNRKSPIRYSITRSKVFNINDNGTITTSNSLDREISA
WYNLSITATEKYNIEQISSIPLYVQVLNINDHAPEFSQYYETYVCENAGSGQVIQTISAVDR
DESIEEHHFYFNLSVEDTNSSFTIIDNQDNTAVILTNRTGFNLQEEPVFYISILIADNGIP
SLTSTNTLTIHVCDGDSGSTQTCQYQELVLSMGFKTEVIIAILICIMIIFGFIFLTLGLKQ
RRKQILFPEKSEDFRENIFQYDDEGGGEEDTEAFDIAELRSSTIMRERKTRKTTSAEIRSLY
RQSLQVGPDSAIFRKFILEKLEEANTDPCAPPFDSLQTYAFEGTGLAGSLSSLESASVSDQD
ESYDYLNELGPRFKRLACMFGSAVQSN

Important features:

Signal peptide:

amino acids 1-21

Transmembrane domain:

amino acids 597-617

N-glycosylation sites.

amino acids 57-60, 74-77, 419-423, 437-440, 508-511, 515-518,
516-519 and 534-537

Cadherins extracellular repeated domain signature.

amino acids 136-146 and 244-254

0978403-101601

FIGURE 96

ATTTCAAGGCCAGCCATATTTTTNTGTTGAACCAACAACAGGAGTCATAAGAATATTTNTA
AAATGGATAGAGAACTGCAAGATGAGTATTGGGTAATCATTCAAGCCAAGGACATGATTGGT
CAGCCAGGAGCGTTGTNTGGAACAACAAGTGTATTAATTAACTTTCAGATGTTAATGACAA
TAAGCCTATATTTAAAGAAAGTTTATACCGCTTGACTGTNTNTGAATCTGCACCCACTGGGA
NTTNTATAGGAACAATCATGGCATATGATAATGACATAGGAGAGAATGCAGAAATGGATTAC
AGCATTGAAGAGGATGATTGCGAAACATTTGACATTATT

09373403.101601

FIGURE 97

GCAACCTCAGCTTCTAGTATCCAGACTCCAGCGCCGCCCCGGGCGCGGACCCCAACCCCGAC
CCAGAGCTTCTCCAGCGGGCGCGCAGCGAGGGCTCCCCGCTTAACTTCTCCGCGGGG
CCCAGCCACCTTCGGGAGTCCGGGTGCCCCACCTGCAAACCTCTCCGCTTCTGCACCTGCCA
CCCCTGAGCCAGCGCGGGCCCCCGAGCGAGTCA**ATGGCC**AAACGCGGGGCTGCAGCTGTTGGGC
TTATTCTCGCCTTCTCTGGGATGGATCGGCGCCATCGTCAGCACTGCCCTGCCCGAGTGGAG
GATTTACTCCTATGCCGGCGACAACATCGTGACCGCCAGGCCATGTACGAGGGGCTGTGGA
TGTCTGCGTGTGCGAGAGCACCGGGCAGATCCAGTGCAAAGTCTTTGACTCCTTGCTGAAT
CTGAGCAGCACATTGCAAGCAACCCGTGCCCTTGATGGTGGTTGGCATCCTCCTGGGAGTGAT
AGCAATCTTTGTGGCCACCGTTGGCATGAAGTGATGAAGTGCTTGGAAAGACGATGAGGTGC
AGAAGATGAGGATGGCTGTCTATTGGGGGTGCGATATTTCTTCTTGCAAGTCTGGCTATTTTA
GTTGCCACAGCATGGTATGGCAATAGAATCGTTCAGAATCTATGACCCCTATGACCCCACT
CAATGCCAGGTACGAATTTGGTCAGGCTCTCTTCACTGGCTGGGCTGCTGCTTCTCTCTGCC
TTCGGGAGGTGCCCTACTTTGCTGTTCTGTCCCCGAAAAACAACCTCTTACCCAACACCA
AGGCCCTATCCAAAACTGCACCTTCAGCGGGAAAGACTACGTG**TGA**CACAGAGGCAAAAG
GAGAAAATCATGTTGAAACAAACCGAAAATGGACATTGAGATACTATCATTAACTAGGAC
CTTAGAATTTGGGTATTGTAATCTGAAGTATGGTATTACAAAACAAACAAACAAAA
ACCCATGTGTTAAATACTCAGTGTAAACATGGCTTAATCTTATTTTATCTTCTTCTCTCA
ATATAGGAGGGAAGATTTTTCCATTGTTACTGCTTCCCATGAGTAATCATACTCAAAT
GGGGGAAGGGGTGCTCCTTAAATATATAGATATGATATACATGATTTTTCTATTTAAAA
ATAGACAGTAAATACTATTCTCATTATGTTGATACTAGCATACTTAAATATCTCTAAAT
AGGTAAATGTATTTAATTCATATTGATGAAGATGTTTATTGGTATATTTCTTTTTCGTCC
CATATACATATGTAACAGTCAAATGCATTACTTCTTCTTCAATGCTTTGGGTGCTTTG
CCACAAGACCTAGCCTAATTTACCAAGGATGAATCTTTCAATTCTTCATGCGTGCCCTTTT
CATATACTTATTTTATTTTTTACCATAATCTTATAGCACTTGCACTGCTTATTAAGCCCTTAT
TTGTTTTGTGTTTCAATTGGTCTCTATCTCCTGAATCTAACACATTTATAGCCTACATTTTA
GTTTCTAAAGCCAAGAAGATTTATTACAAATCAGAACTTTGGAGGCAAATCTTCTGCATG
ACCAAAGTGATAAATTCCTGTTGACCTTCCACACAATCCCTGTACTCTGACCCATAGCACT
CTTGTTTGCTTTGAAAATATTTGTCCAATTGAGTAGCTGCATGCTGTTCCCCAGGTGTTGT
AACCAACTTTATTGATTGAATTTTTAAGCTACTTATTATAGTTTTTATATCCCCTAAACT
ACCTTTTTGTTCCTTCCCTTAATTTGTATGTTTTCCCAAGTGTAATTATCATGCGTTTTTA
TATCTTCTAATAAGGTGTGGTCTGTTTGTCTGAACAAAGTGCTAGACTTTCTGGAGTGATA
ATCTGGTGACAAATATCTCTCTGTAGCTGTAAGCAAGTCACTTAATCTTCTACCTTTTTT
TTCATCTGCCAAATTGAGATAATGATACTTAACCAGTTAGAAGAGGTAGTGTGAATATTAA
TTAGTTTATATTACTCTTATCTTTGAACATGAACATAGCCTATGTAGTGTCTTATTTGCT
CAGTGGCTGAGACACTGAAGAAGTCACTGAACAAAACCTACACACGTACCTTCATGTGATT
CACTGCCTTCTCTCTACCAGTCTATTTCCTACTGAACAAAACCTACACACATACCTTCAT
GTGGTTCACTGTCCTTCTCTCTACACAGTCTATTTCCTACTGAACAAAACCTACGACATAC
CTTCATGTGGCTCAGTGCCTTCTCTCTACCAGTCTATTTCCTACTTTCAGTGTGTCT
GACATGTTTGTGCTCTGTTCTTTTAAACAACTGCTCTTACTTTTCCAGTCTGTACAGAATG
CTATTTCACTTGAGCAAGATGATGTAATGAAAGGGTGTGGCACTGGTGTCTGGAGACCTG
GATTTGAGTCTTGGTGCTATCAATCACCGTCTGTGTTGAGCAAGGCATTGGCTGCTGTAA
GCTTATTGCTTCATCTGTAAGCGGTGGTTTGTAAATTCGTATCTTCCACCTCACAGTGATG
TTGTGGGGATCCAGTGAGATAGAATACATGTAAGTGTGGTTTTGTAAATTTAAAAGTGTAT
ACTAAGGGAAAGAATTGAGGAATTAACGTACATACGTTTGTGGTGTGCTTTCAAATGTTTGA
AAATAAAAAAATGTTAAG

FIGURE 98

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA52185

><subunit 1 of 1, 211 aa, 1 stop

><MW: 22744, pI: 8.51, NX(S/T): 1

MANAGLQLLGFI~~L~~AFLGWIGAIVSTALPQWRIYSYAGDNIVTAQAMYEG~~L~~WMSCVSQSTGQI
QCKVFDSLNL~~S~~STLQATRALMVVGILLGVIAIFVATVGMKCMKCLEDD~~E~~VQKMRMAVIGGA
IFLLAGLAILVATAWYGNRIVQEFYDPMT~~P~~VNARYEFGQALFTGWAAASLCLLGGALLCCSC
PRKTTSYPTPRPYPKPAPSSGKDYV

Important features:

Signal peptide:

amino acids 1-21

Transmembrane domains:

amino acids 82-102, 118-142 and 161-187

N-glycosylation site.

amino acids 72-75

PMP-22 / EMP / MP20 family proteins

amino acids 70-111

ABC-2 type transport system integral membrane protein

amino acids 119-133

FIGURE 99

TTCTGGCCAAACCCGGGGCTNCAGCTGTTGGGCTTCATCTCGCCTTCCTGGGATGGATCGGC
GCCATCNTCACACTGCCCTTCCCCAGTGGAGGATTTTACTCCCTATGCTGGCGACAACATCG
TGACCCGCCAGCCCATGTACGAGGGGCTGTGGATGTCCNGCGTGTGCGAGAGCACCGGGCAG
ATCCAGTGCAAAGTCTTTGACTCCTTGCTGAATCTGAGCAGCACATTGCAAGCAACCCGTGC
CTTGATGGTGGTTGGCATCCTCCTGGGAGTGATAGCAATCTTTGTGGCCACCGTTGGCATGA
AGTGTATGAAGTGCTTGAAGACGATGAGGTGCAGAAGATGAGGATGGCTGTTCATTGGGGGC
GCGATATTTCTTCTTGCAAGTCTGGCTATTTTAGTTGCCACAGCATGGTATGGCAATAGAAN
CNTTCAACANTTCTATGACCCTATGACCCAGTCAATGCCAGGTACGAATTTGGTCA
GGCTCTCTTCACTGGCTGGGCTGCTGCTTCTCTCTGCTTCTGGGAGGTGCCCTACTTTGCT
GTTCTGTGCC

45341 204545

FIGURE 100

ACCCTTGACCCAACGCGGCCCCCGACCGNTTCATGGCCAAACGCGGGNCTCCAGCTGTTGG
GCTTCATTCTCCCCTTCCTGGGATGGACCGGCGCCCATCNTCAGCACTGCCCTGCCCCAGTG
GAGGATTTACTCCTATNCCGGCNACAACATCGTGACCGCCCAGGCCNTGTACGAGGGGCTGT
GGATGTCCTGCGTGTCGACAGCACCGGGCAGATCCAGTGCAAAGTCTTTGACTCCCTTGCT
GAATCTGAGCAGCACATTGCAAGCAACCCGTGCCTTGATGGTGGTTGGCATCCTCCTGGGAG
TGATAGCAATCTTNNTGGCCACCGTTGTNNNTGAAGTGTATGAAGTGCTTGGAAGACGATGA
GGTGCAGAAGATGAGGATGGCTGTTCATTGGGGGCGCGATATTTCTTCTTGCAAGTCTGGCTA
TTTtagTTGCCACAGCATGGTATGGCAATAGAATCGTTCAAGAATTCTATGACCCTATGACCGA

090307

GGGCCCCGACCATTATCCAAACCGGNTCACTGTTGGCTCATCTCCCTCCTGGATGAANCGCGC
CATCNTCAGACTCCCTGCCCATGGAGATTNNCCATGCTGGCGACAACATCNTGACCCCC
AGCCATGTACGAGGGGCTTTGAACGTCNCGTGTGCGAGANACCGGGCAGATCCAGTGCAA
AGTCTTTGACTCCTTGCTGAATCTGNGCAGCAATTGCAGCAACCNTGCCCTGATGGTGGT
TGGCATCCTCCTGGGAGTGATAGCAATCTTTGTGGCCACCGTTGGCATGAAGTGATGAAGT
GCTTGAAGACGATGAGGTGCGAAGATGAGGATGGCTGTCAATTGGGGGCGCGATATTTCCT
CTTGCAAGTCTGGCTATTNNNGTTGCCACAGCATGGTATGGCAATAGAATCGTTCAGAAT
TCTATGACCCATGACCCGAGTCAATGCCAGGTACGAATTTGGTCAAGGCTCTCTTCACTGGC
TGGGCTGCTGCTTCTCTGCCTTCTGGGAGGTGCCCTACTTTGCTGTTCTCTGCGA

FIGURE 102

ATTCTCCCCCTCCTGGATGGATCGCNCCACCGTCACATTGCCTTCCCCANTGGAGGATTNAC
TCCTATGCTGGCGACAACATCGTGACCCCCCAGGCCATTACCGAGGGGCTTTGGATGTCNT
GCNTGTCGCAGAGCACCGGGCAGATCCCAGTGCAAAGTCTTTGACTCCTTGCTGAATCTGAG
CAGCACATTGCAAGCAACCCGTGCCTTGATGGGGTTGGCATCCTCCTGGGAGTGATAGCAAC
CTTTGTGGCCACCGTTGGCATGAAGTGTATGAAGTGCTTGAAGACGATGAGGTGCCAGAAG
ATGAGGATGGCTGTCAATTGGGGGCGCGATATTTCTTGTTCAGGTCTGGCTATTTTAGTNGC
CACAGCATGGTATGGCAATAGANTNNTTCNNGNNNTCTATGACCCCTATGACCCCAAGTCAATG
CCAGGTACGAATTTGGTCAGGCTCTCTTCACTGGCTGGGCTGCTGCTTCTCTCTGCCTTCTG
GGAGGTGCCCTACTTTGCTGTTCTCTGTCCC

109104-2402-560

FIGURE 103

AGAGCACCGGCAGATCCCAGTNCAAAGTCTTTGACCCCTTGCTGAATCTGAGCAGCACATTNC
AAGCAACCCCTTGCCCTTGAAGGTGGTTGNCATCCCCCTGGGAGTGAATAGCAATCTTTGTG
GCCACCGTTGGCATGAAGTNTATGAAGTGCTTGGAAGACGATGAGGTGCAGAAGATGAGGAT
GGCTGTCATTGGGGGCGCGATATTTCTTCTTGACAGGTCTGGCTATTTTAGTNNCCACAGCAT
GGTATGGCAATAGNATNNITTCGNGGNTTCTATGACCCCTATGACCCCAAGTCAATGCCAGGTAC
GAATTTGGTCAGGCTCTCTTCACTGGCTGGGCTGCTGCTTCTCTCTGCCTTCTGGGAGGTGC
CCTACTTTGCTGTTCTCTGTCCCCGAA

FIGURE 104

AGCAATGCCCTGCCCCAGTGGAGGATTAATTCCTATGNTGGGGACAACATTGTGACNGCCC
AGGCCATGTACGGGGGGCTGTGGATGTCCTGCGTGTGCGAGAGCACCGGGCAGATCCAGTGC
AAAGTNTTTGACTCCTTGCTGAATTTGAGCAGCACATTGCAAGCAACCCGTGCCTTGATGGT
GGTTGGCATCTTCCTGGGAGTGATAGCAATCTTTGTGGCCACCGTGGNAATGAAGTGTATGA
AGTGCTTGGAAGACGATGAGGTGCAGAAGATGAGGATGGCTGTCATTGGGGGCGCGATATTT
CTTNTTGCAGGTCTGGCTATTTTAGTTGCCACAGCATGGTATGGCAATAGAATNGTTC AAGA
ATTTTATGACCCATGACCCCAGTCAATGCCAGGTACGAATTTGGTCAGGCTTTNTTCACTG
GCTGGGCTGCTGCTTNTTCTGCCTTNTGGGAGGTGCCCTANTTTGCTGTTCTGCGAACC

109101-2042500

FIGURE 105

TCATAGGGGGCGCGATATTTTTTCTTGCAAGTNTGGTTATTTTAGTTGCCACAGCATGGTA
TGGCAATAGAATCGTTCAAGAATTNTATGACCCATGACCCAGTCAATGCCAGGTACGAAT
TTGGTCAGGCTCTNTTCACTGGNTGGGCTGCTGCTTCTNTNNGCCTTNTGGGAGGTGCCCTA
CTTTGCTGTTCCCTG

10501-105260

FIGURE 106

TTCTCTGGGATGGATCCGCCCCCATCNTCACATGCCCTGCCCCNTGGAGATTTACNCCTATGC
TGGCGAACAAACATCNTGACCGCCCAGGCCATGTACGAGGGGCTGTGGAATGTCCTGCCGTGTC
CCAGAGCACCGGGCAGATCCAGTGCAAAGTCTTTGACTCCTTGCTGAATCTGAGCAGCACAT
TGCAAGCAACCNTGCCTTGATGGTGGTTGGCATCCTCCTGGGAGTGATAGCAATCTTTGTGG
CCACCGTTGGCATGAAAGTGTATGAAGTGCTTGAAGACGATGAGGTGCAGAAGATGAGGAT
GGCTGTCAITGGGGGCGCGATATTTCTTCTTGCAAGTCTGGCTATTTTAGNNGCCACAGCAT
GGTATGGCAATCAGACCCNNTCANAACTCTATGACCCTATGACCCAGTCAATGCCAGGTA
CGAATTTGGTCAGGCTCTCTTCACTGGCTGGGCTGCTGCTTCTCTCTGCCTTCTGGGAGGTG
CCCTACTTTGCTGTTCTGTCCCCGAAAAACAACCTCTTACCCACG

057843-10450

FIGURE 107

CGGGGCTGCAGCTGTTGGGCTTCATCTCGCTTCCTGGGATGGAATCGGCGCCATCGTCAGCA
CTGCCCCTGCCCCATGGAGGATTTACTCNTATGCTGGCGACAACATCGTGACCNCCCAGGCCA
TGACGAGGGGCTGTGGATGTCNGCGTGTGCGCAGACACCGGGCAGATCCAGTGCAAAGTCT
TTGACTCCTTGCTGAATCTGAGCAGCACATTGCAAGCAACNTGCCTTGATGGTGGTTGGCA
TCCTCCTGGGAGTGATAGCAATCTTTGTGGCCACCGTTGGCATGAAGTGTATGAAGTGCTTG
GAAGACGATGAGGTGCAGAAGATGAGGATGGCTGTCATTGGGGGCGCGATATTTCTTCTTG
AGGTCTGGCTATTTNTAGTTGCCACAGCATGGTATGGCAATAGAATCGTTCAAGAATTCTAT
GACCCATGACCCAGTCAATGCCAGGTACGAATTTGGTCAGGCTCTCTTCACTGGCTGGGC
TGCTGCTTCTCTGCTTCTGGGAGGTGCCCTACTTTGCTGTTCTGCGAA

FIGURE 108

GCGTGCCGTCAGCTCGCCGGGACCGCGGCCCTCGCCCTCGCCCTCGCCCTCGCCCTGCGCCTGCAC
CGCGTAGACGACCCCCCTCCAGCGCGCCACCCGGTAGAGGACCCCGCCCGTGCCCG
ACCGGTCCCCGCTTTTTTGTAAACTTAAAGCGGCGCAGCATTAAAGCTTCCCGCCCCGGT
GACCTCTCAGGGGTCTCCCCGCAAGGTGCTCGCGCGCTAAGGAACATGGCGAAGGTGGAG
CAGGTCTGAGCCTCGAGCCGACGACGAGCTCAAATTCGAGGTCCCTTCACCGATGTTGT
CACCACCAACCTAAAGCTTGGCAACCCGACAGACCGAAATGTGTGTTTTAAGGTGAAGACTA
CAGCACCACGTAGGTACTGTGTGAGGCCCAACAGCGGAATCATCGATGCAGGGGCCCTCAATT
AATGTATCTGTGATGTTACAGCCTTTCGATTATGATCCCAATGAGAAAAGTAAACACAAGTT
TATGGTTTCAGTCTATGTTTGTCTCCAACCTGACACTTCAGATATGGAAGCAGTATGGAAGGAG
CAAAACCGGAAGACCTTATGGATTCAAACCTTAGATGTGTGTTGAATTGCCAGCAGAGAAT
GATAAACCACATGATGTAGAAATAAATAAAATTTATATCCACAACCTGCATCAAAGACAGAAAC
ACCAATAGTGTCTAAGTCTCTGAGTCTCTTCTTTGGATGACACCGAAGTTAAGAAGGTTATGG
AAGAATGTAAGAGGCTGCAAGGTGAAGTTGAGAGGCTACGGGAGGAGAACAGCAGTTCAAG
GAAGAAGATGGA CTGCGGATGAGGAAGACAGTGCAGAGCAACAGCCCCATTTGAGCATTAGC
CCCAACTGGGAAGGAAGAGCCTTAGCACCCGGCTCTTGGCTCTGGTGTTTGTGTTCTTTA
TCGTTGGTGTAAATTTATGGGAAGATTGCCTTGTAGAGGTAGCATGCACAGGATGGTAAATTG
GATTGGTGGATCCACCATATCATGGGATTTAAATTTATCATAACCATGTGTAAAAAGAAATT
AATGTATGATGACATCTCACAGGTCTTGCCCTTTAAATTACCCCTCCCTGCACACACATACAC
AGATACACACACAAATATAATGTAACGATCTTTTAGAAAAGTTAAAAATGTATAGTAACTG
ATTGAGGGGGAAAAAGAATGATCTTTATTAATGACAAGGGAACCATGAGTAATGCCACAAT
GGCATATTTGTAATGTCAATTTAAACATTTGGTAGGCCTTGGTACATGATGCTGGATTACCTC
TCTTAAATGACACCCCTTCTCGCCTGTGGTGTCTGGCCCTTGGGGAGCTGGAGCCCAGCAT
GCTGGGGAGTGGGTGAGCTCCACACAGTAGTCCCCACGTGGCCCACTCCCGGCCAGGCTG
CTTTCCGTGTCTTCAGTTCTGTCCAAAGCCATCAGCTCCTTGGGACTGATGAACAGAGTCAGA
AGCCCAAAGGAATTGCACTGTGGCAGCATCAGACGTACTCGTCATAAGTGAGAGGCGTGTGT
TGACTGATTGACCCAGCGCTTTGGAAATAAATGCGAGTGCTTTGTTCACCTTAAAGGGACCAA
GCTAAATTTGTATTGGTTTATGTAGTGAAGTCAAACCTGTTATTTCAGAGATGTTTAAATGCATA
TTTAACTTATTTAATGTATTTTCATCTCATGTTTTCTTATTGTCAACAAGGTACAGTTAATGC
TGCGTGTCTGCTGAACCTGTTTGGGTGAACCTGGTATTGTGCTGGAGGGCTGTGGGCTCCTCT
GTCTCTGGAGAGTCTGGTCATGTGGAGGTGGGGTTTATTTGGGATGCTGGAGAAGAGCTGCCA
GGAAGTGTTTTTCTGGGTGAGTAAATAACAACCTGTCTAGGGGAGGAAATTTCTCAGTAGTG
ACAGTCAACTCTAGGTTACCTTTTTTAATGAAGAGTAGTCAGTCTCTTAGATTGTTCTTTATA
CCACCTCTCAACCATTACTCACACTTCCAGCGCCAGGTCCAAGTCTGAGCCTGACCTCCCC
TTGGGACCTAGCCTGGAGTCAGGCAAAATGGATCGGGCTGCAGAGGGTTAGAAGCGAGGGC
ACCAGCAGTTGTGGTGGGGAGCAAGGGAAGAGAGAACTCTTCAGCGAATCCTTCTAGTAC
TAGTTGAGAGTTTGACTGTGAATTAATTTTATGCCATAAAAGACCAACCCAGTTCTGTTTTGA
CTATGTAGCATCTTGAAAAGAAAAATTTATAATAAAGCCCCAAATTAAGAAAA

FIGURE 109

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA53977

<subunit 1 of 1, 243 aa, 1 stop

<MW: 27228, pI: 7.43, NX(S/T): 2

MAKVEQVLSLEPQHCLKFRGPFTDVVTTNLKLGNPTRNVCFKVKTTAPRRYCVRPNSGIID
AGASINVSVMQLQPFYDPNEKSKHKFMVQSMFAPTDTSMEAVWKEAKPEDLMDSKLRVFE
LPAENDKPHDVEINKIIISTASKTETPIVSKSLSSSLDDTEVKKVMEECKRLQGEVQRLREE
NKQFKEEDGLMRKTVQSNPISALAPTGKEEGLSTRLLALVVLFFIVGVIIIGKIAL

Important features:

Transmembrane domain:

amino acids 224-239

N-glycosylation site.

amino acids 68-71

N-myristoylation site.

amino acids 59-64, 64-69 and 235-240

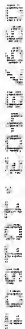


FIGURE 110

GTCAGTCTTCTAGATTGTCCTTATCCCACCTTTCAACCANTACTCACATTTNAGCGCCCAG
GTCCANGTCTGAGCCTGACTTCCCCCTTGGGGACCTAGCCTGGAGTCAGGACAATGGNTCGGG
CTGCAGAGGNTTAGAAGCGAGGGCACCAGCAGTTTGGGTGGGGAGCAAGGGNNGAGAGAAA
CTCTTCAGCGAATCCTTCTAGTACTAGTTGAGAGTTTGACTGTGAATTAAATTTATGCCATA
AAAGACNAACCCAGTTCTGTTTGACTATGTAGCATCTTGAAAAGAAAAATTATAATAAAGCC
CCAAAATTAAGAATTCTTTTGTCAATTTGTACATTTGCTCTATGGGGGAATTATTATTTT
ATCATTTTTATTATTTTGCCATTGGAAGGTTAACTTTAAATGAGC

FIGURE 111

TATTGTAAAGGCCATTTTAAACCATTGGTAGGCCCTTGGTACATGATGCTGGATTACCTCCTT
AAATGACACCNTTCCTCGCCTGTTGGTGCTGGCCNTTGGGGAGCTGGAGCCCCAGCATGCTG
GGGAGTGCGGTCAGCTCCACACAGTAGTCCCCACGTGGCCCCACTCCCGGGCCAGGCTGCTTT
CCGTGTCTTCAGTTCTGTCCAAGCCATCAGCTCCTTGGGACTGATGAACAGAGTCAGAAGCC
CAAAGGAATTGCCACTGTGGCAGCATCAGACGTA CT CGT CATAAGTGAGAGGCGTGTGTGA
CTGATTGACCCAGCGCTTTGGAAATAAATGGCAGTGCTTTGTTCACTTAAAGGGACCAAGCT
AAATTGTATTGGTTCATGTAGTGAAGTCAAAC T GTTATTCAGAGATGTTTAATGCATATTTA
ACTTATTTAATGTATTT CATCTCATGTTTTCTTATTGTCACAAGAGTACAGTTAATGCTGCG
TGCTGCTGAACTCTGTTGGGTGAACTGGTATTGCTGCTGGAGGGCTG

CCCTGGTGGTTTTGTTCTTTAATTGTTGGTGTAATTNTTGGGAAGATTGCTTGTAGAGGTA
GNATGCACCNNGGCTGGTAAATTGGATTGGTGGATCCACCATATCCATGGGATTTAAATTTAT
CATAACCATGTGTAAAAAGAAATTAATGTATGATGACATNTCAGAGTATTGCCTTTAAATT
ACCCATCCCTGNANACACATACACAGATACACANANACAAATNTAATGTAAACGATNTTTTAG
AAAGTTAAAAATGTATAGTAAC

FIGURE 113

GGTGGCCCATTCCTGGCCAGGCTGCTTTCCGGTNTTCAGTTCTGTCCAAGCCATCAGCTCC
TTGGGACTGATGAACAGAGTCAGAAGCCCAAAGGAATTGCACTGTGGCAGCATNAGACGTAC
TTGTNATAAGTGAGAGGCGTGTGTTGACTGATTGACCCAGCGCTTTGGAAATAAATGGCAGT
GCTTTGTTTCANTTAAAGGGACCAAGCTAAATTTGTATTGGTTCATGTAGTGAAGTCAAACCTG
TTATTCAGAGATGTTTAATGCATATTTAANTTATTTAATGTATTTNAINTCATGTTTTCCTTA
TTGTCACAAGAGTACAGTTAATGCTGCGTGCTGCTGAANTNTGTTGGGTGAACTGGTATTGC
TGCTGGAGGGCTGTGGGCTCCTCTGTCTTTGGAGAGTCTGGTCATGTGGAGGTGGG

109101: 00000000

Figure 1 is a schematic representation of the experimental design. It shows a sequence of steps: 1. Pre-test (N=10), 2. Training (N=10), 3. Transfer (N=10), 4. Post-transfer (N=10), 5. Post-transfer (N=10), 6. Post-transfer (N=10), 7. Post-transfer (N=10), 8. Post-transfer (N=10), 9. Post-transfer (N=10), 10. Post-transfer (N=10). Each step includes a small schematic of the experimental setup.

FIGURE 115

AAACCTTTAAAAGTTGAGGGGAAAAGAATGATCCTTTATTAATGACAAGGGAAACCNTGNGT
AATGCCACAATGGCATATTGTAAATGTCATTTTAAACATTGGTAGGCCTTGGTACATGATGC
TGGATTACCTCTCTTAAATGACACCCCTTCCTCGCCTGTTGGTGCTGGCCCTTGGGGAGCTN
GAGCCCAGCATGCTGGGGAGTGCGGTCTGCTCCACACAGTAGTCCCCANGTGGCCANTCCC
GGCCAGGCTGCTTTCGCTGTCTTCAGTTCGTCCAAGCCATCAGCTCCTTGGGANTGATGA
ACAGAGTCAGAAGCCCCAAAGGAATTGCANTGTGGCAGCATCAGANGTANTNGTCATAAGTGA
GAGGCGTGTGTTGANTGATTGACCCAGCGCTTTGGAATAAATGGCAGTGCTTTGTTTCAANT
AAAGGGNCCAAGNTAAATTTGTATTGGTTCATGTAGTGAAGTCAAANTGTTATTTCAGAGATG
TTTAATGCATATTTAANTTATTTAATGTATTTCAATNTCATGTTTCTTATTGTCACAAGGGT
ACAGTTAATGCTGCGTGCTGCTGAANTCTGTTGGGTGAANTGGTATTGCTG

0978403-101601

JOSEPH

GGCCCTTGGGGAGCTGGAGCCAGCATGCTGGGGAGTGCGGTCAGCTCCACACAGTAGTCCC
CAGTGGCCCACTCCCGGCCAGGCTGCTTTCGTGTCTTCAGTTCTGTCCAAGCCATCAGC
TCCTTGGGACTGATGAACAGAGTCAGAAGCCCAAAGGAATTGCACTGTGGCAGCATCAGACG
TACTCGTCATAAGTGAGAGGCGTGTGTTGACTGATTGACCAGCGCTTTGGAAATAAATGGC
AGTGCTTTGTTCACTTAAAGGGACCAAGCTAAATTGTATTGGTTCATGTAGTGAAGTCAA
CTGTTATTACAGAGATGTTTAAATGCATATTTAACTTATTTAATGTATTTTCATCTCATGTTTTC
TTATTGTACAAAGAGTACAGTTAATGCTGCGTGCTGCTGAACTGTTGGGTGAACTGGTAT
TGCTGCTGGAGGGCTGTGGGCTCCTCTGCTCTGAGAGTCTGGTCATGTGGAGGTGGG

FIGURE 117

GCGAGCTCCGGGTGCTGTGGCCCCGGCCTTGGCGGGGCGGCCTCCGGCTCAGGCTGGCTGAGA
GGCTCCACGCTGCAGCGTCCCCGCCGCCCTCCTCGGGAGCTCTGATCTCAGCTGACAGTGCC
CTCGGGGACCAAAACAAGCCTGGCAGGGTCTCACTTTGTTGCCAGGCTGGAGTTCACTGCCA
TGATCATGGTTTACTGACGCCTTGACCTCCTGGGTCAAGCGATCCTGCTGAGTAGCTGGGA
CTACAGGACAAAAATTAGAAGATCAAAATGGAAAAATATGCTGCTTTGGTTGATATTTTTCACC
CCTGGGTGGACCTTCATTGATGGATCTGAAATGGAATGGGATTTTATGTGGCACTTGAGAAA
GGTACCCCGGATTGTGAGTAAAGGACTTTCCATCTCACCAGCCCCGATTGTGAGGCAGATG
CTAAGATGATGGTAAATACAGTGTGTGGCATCGAATGCCAGAAAGAACTCCCAACTCCCAGC
CTTCTGAATTGGAGGATTATCTTTCCATGAGACTGTCTTTGAGAATGGCACC CGAACCTT
AACCAGGGTGAAAGTTCAAGATTTGGTTCTTGAGCCGACTCAAATATCACCACAAAGGGAG
TATCTGTTAGGAGAAAGAGACAGGTGTATGGCACCAGACAGCAGGTTCAAGCATCTTGGACAAA
AGGTTCTTAACCAATTTCCCTTTCAGCACAGCTGTGAAGCTTCCACGGGCTGTAGTGGCAT
TCTCATTTCCTCAGCATGTTCTAACTGCTGCCACTGTGTTTATGATGGAAGGACTATG
TCAAAGGGAGTAAAAAGCTAAGGGTAGGGTGTGTTGAAGATGAGGAATAAAAGTGGAGGCAAG
AAACGTGAGGTTCTAAGAGGAGCAGGAGAGAAGCTAGTGGTGGTGACCAAGAGAGGGTAC
CAGAGAGCATCTGCAGGAGAGAGCGAAGGGTGGGAGAAGAAGAAAAAATCTGCGCCGGGTC
AGAGGATTGCCAAGGGAGGCCCTTCTTTTCACTGAGCCCGGTTCAAGAAATACCCACATTCCG
AAGGGCTGGGCACGAGGAGGATGGGGGACGCTACCTTGGACTATGACTATGCTCTTCTGGA
GCTGAAGCGTGCTCACAAAAAGAAAATACATGGAACCTGGAATCAGGCCAACGATCAAGAAAA
TGCCTGGTGAATGATCCACTTCTCAGGATTTGATAACGATAGGGCTGATCAGTTGGTCTAT
CGGTTTTGCAGTGTGTCGACGAATCCAATGATCTCCTTTACCAATACTGCGATGCTGAGTC
GGGCTCCACCGGTTCCGGGGTCTATCTGCGTCTGAAAGATCCAGACAAAAAGAATTGGAAGC
GCAAAATCATTTGCGGTCTACTCAGGGCACCAGTGGGTGGATGTCCACGGGGTTCAAGAGGAC
TACAACGTTGCTGTTTCGCATCACTCCCTTAAAAATACGCCCAGATTTGCCTCTGGATTACGG
GAACGATGCCAATTGTGCTTACGGCTAAACAGAGACCTGAAACAGGGCGGTGTATCATCTAAA
TCACAGAGAAAAACAGCTCTGCTTACCGTAGTGAGATCACTTCATAGTTATGCTTGGACTT
GAACTCTGTCATAGCATTTCAACATTTTTCAAATCAGGAGATTTTCGTCCATTTAAAAAA
TGTATAGGTGCAGATATTGAAACTAGGTGGGCACTTCAATGCCAAGTATATACCTCTCTTTA
CATGGTGATGAGTTTCATTGTTAGAAAAATTTTGTGCTTCTTAAAAATTAGACACACTTT
AAACCTTCAAACAGGTATTATAAATAACATGTGACTCCTTAATGGACTTATTCTCAGGGTCC
TACTCTAAGAAGATCTAATAGGATGCTGGTTGTGTATTAAATGTGAAATTGCATAGATAAA
GGTAGATGGTAAAGCAATTAGTATCAGAATAGAGACAGAAAGTTCAACACACAGTTTGACTA
CTCTGAGATGGATCCATTCACTCATGCCCTCAATGTTTATATTGTGTTATCTGTTGGGTCT
GGGACATTTAGTTTAGTTTTTTTGAAGAATTACAAATCAGAAAGAAAAGCAAGCATTATAAA
CAAACTAATAACTGTTTTACTGCTTTAAGAAATAACAATTACAATGTGTATTATTTAAAAA
TGGGAGAAAATAGTTTTGTTCTATGAAATAAACCTAGTTTGAATAAGGGAAGCTGAGACATTT
TAAGATCTCAAGTTTTTATTTAACTAATACTCAAATATGGACTTTTATGTATGCATAGGG
AAGACACTTCACAAATTATGAATGATCATGTGTTGAAAGCCACATTATTTTATGCTATACAT
TCTATGTATGAGGTGCTACATTTTAGGACAAAGAATTCTGTAATCTTTTCAAGAAAGAGT
CTTTTCTCCTTGACAAAATCCAGCTTTTGTATGAGGACTATAGGGTGAATTTCTGATTAG
TAATTTTAGATATGCTCTTCTTAAAAATGAATAAAATTTATGAATATGA

00072403-30600

FIGURE 118

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA57253

<subunit 1 of 1, 413 aa, 1 stop

<MW: 47070, pI: 9.92, NX(S/T): 3

MENMLLWLIFFTPGWTLIDGSEMEWDFMWHLRKVPRIVSERTFHLTSPAFEADAKMMVNTVC
GIECQKELPTPSLSELEDYLSYETVFENGTRTLTRVKVQDLVLEPTQNITTKGVSVRKRQV
YGTDSRFSILDKRFLTNPFPSTAVKLSTGCSGILISPQHVLTAACHVDGKDYVKGSKKLRV
GLLKMRNKSGGKKRRGSKRSRREASGGDQREGTREHLQERAKGRRRKSGRGQRIAEGRPS
FQWTRVKNTHIPKGWARGMGDATLDYDYALLELKRAHKKKYMELGISPTIKMPGGMIHFS
GFDNDRADQLVYRFCSVSDESNDLLYQYCDAESGSTGSGVYLRLKDPDKKNWKRKIIAVYSG
HQWVDVHGVQKDYNAVAVRITPLKYAQICLWIHGNDANCAYG

Important features:

Signal peptide:

amino acids 1-16

N-glycosylation sites.

amino acids 90-93, 110-113 and 193-196

Glycosaminoglycan attachment site.

amino acids 236-239

Serine proteases, trypsin family, histidine active site.

amino acids 165-170

FIGURE 119

AATGTGAGAGGGGCTGATGGAAGCTGATAGGCAGGACTGGAGTGTAGCACCAGTACTGGAT
GTGACAGCAGGCAGAGGAGCACTTAGCAGCTTATTTCAGTGTCCGATTCTGATTCCGGCAAGG
ATCCAAGCATGGAATGCTGCCGTCGGGGCAACTCCTGGCACACTGCTCCTCTTTCTGGCTTTC
CTGCTCCTGAGTTCAGGACCGCACGCTCCGAGGAGGACCGGGACGGCCTATGGGATGCCTG
GGGCCCATGGAGTGAATGCTCACGCACCTGCGGGGGAGGGGCCTCTACTCTCTGAGGCGCT
GCCTGAGCAGCAAGAGCTGTGAAGGAAGAAATATCCGATACAGAACATGCAGTAATGTGGAC
TGCCACCAGAAGCAGGTGATTTCCGAGCTCAGCAATGCTCAGCTCATAATGATGTCAAGCA
CCATGGCCAGTTTTATGAATGGCTTCTGTGTCTAATGACCTGACAACCCATGTTCACTCA
AGTGCCAAGCCAAAGGAACAACCTGGTTGTTGAACTAGCACCTAAGGTCTTAGATGGTACG
CGTTGCTATACAGAATCTTTGGATATGTGCATCAGTGGTTTATGCCAAATGTTGGCTGCGA
TCACCAGCTGGGAAGCACCGTCAAGGAAGATAACTGTGGGGTCTGCAACGGAGATGGGTCCA
CCTGCCGGCTGGTCCGAGGGCAGTATAAATCCAGCTCTCCGCAACCAAATCGGATGATACT
GTGGTTGCACTTCCCTATGGAAGTAGACATATTGCGCTTGTCTTAAAGGTCTGTATCACTT
ATATCTGGAACCAAACCTCCAGGGGACTAAAGGTGAAAACAGTCTCAGTCCACAGGAA
CTTTCCTTGTGGACAATTCTAGTGTGGACTTCAGAAATTTCCAGACAAAGAGATACTGAGA
ATGGCTGGACCACTCACAGCAGATTTTATTGTCAAGATTCTGTAACCTCGGGCTCCGCTGACAG
TACAGTCCAGTTCATCTTCTATCAACCCATCATCCACCGATGGAGGGAGACGGATTTCTTTC
CTTGCTCAGCAACCTGTGGAGGAGTTATCAGCTGACATCGGCTGAGTGTCTACGATCTGAGG
AGCAACCGTGTGGTTGCTGACCAATACTGTCACTATTACCCAGAGAACATCAAACCCAAACC
CAAGCTTCAGGAGTGCAACTTGGATCCTTGTCCAGCCAGTGACGGATACAAGCAGATCATGC
CTTATGACCTCTACCATCCCTTCTCTCGGTGGGAGGCCACCCATGGACCGCGTGTCTCTCC
TCGTGTGGGGGGGCATCCAGAGCCGGGCAGTTTCTGTGTGGAGGAGGACATCCAGGGGCA
TGTCACTTCAGTGGAAGAGTGGAATGCATGTACACCCCTAAGATGCCCATCGCGCAGCCCT
GCAACATTTTGA CTGCCCTAAATGGCTGGCACAGGAGTGGTCTCCGTGCACAGTGACATGT
GGCCAGGGCCTCAGATACCGTGTGGTCTCTGCATCGACCATCGAGGAATGCACACAGGAGG
CTGTAGCCCAAAAACAAAGCCCCACATAAAAGAGGAATGCATCGTACCCACTCCCTGTCTATA
AACCCAAAGAGAACTTCAGTTCGAGGCCAAGTTGCCATGGTTCAAACAAGCTCAAGAGCTA
GAAGAAGGAGCTGCTGTGTCTGAGGAGCCCTCGTAAGTTGTAAGACACAGACTGTTCTATA
TTTGAACTGTTTTGTTTAAAGAAAGCAGTGTCTCACTGGTTGTAGCTTTCATGGGTTCTGA
ACTAAGTGTAAATCATCTCACCAAAGCTTTTTGGCTCTCAAATTAAGATTGATTAGTTTCAA
AAAAAAAA

FIGURE 120

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA58847

<subunit 1 of 1, 525 aa, 1 stop

<MW: 58416, pI: 6.62, NX(S/T): 1

MECCRRATPGTLLLFLAFLLLSSRTARSEEDRDGLWDAGPWSECSRTC GGGASYS LRRLCLS
SKSCEGRNIRYRTC SNVDCPPEAGDFRAQQCSAHNDVKHHGQFYEWLPVSNPDNP CSLKQC
AKGTTLVVELAPKVL DGRCTYESLDMCISGLCQIVGCDHQLGSTVKEDNCGVCNGDGSTCR
LVRGQYKSQLSATKSDDTVVALPYGSRHIRLV LKGPDHLYLET KT LQGTKGENSL SSTGTFL
VDNSSVDFQKFDPDKEILRMAGPLTADFIVKIRNSGSADSTVQFIFYQPIIHRWRETDFFP CS
ATCGGGYQLTSAECYDLRSNRVVADQYCHYYPENIKPKPKLQECNLDP CPASDGYKQIMPYD
LYHPLPRWEATPWTACSSSCGGGIQSRAVSCVEEDIQGHVTSVEEWKCMYTPKMPIAQPCNI
FDCPKWLAQEWSPCTVTTCGQGLRYRVVLCIDHRGMHTGGCSPKTKPHIKEECIVPTPCYKPK
EKLPVEAKLPWFKQAQEEGA AVSEEPS

Important features:

Signal peptide:

amino acids 1-25

N-glycosylation site.

amino acids 251-254

Thrombospondin 1

amino acids 385-399

von Willebrand factor type C domain proteins

amino acids 385-399, 445-459 and 42-56

10973437460

FIGURE 121

CGGACGCGTGGGCGGCGGCTGCGGAACTCCCGTGGAGGGGCCGGTGGGCCCTCGGGCCTGAC
AGATGGCAGTGGCCACTGCGGCGGCAGTACTGGCCGCTCTGGGCGGGGCGCTGTGGCTGGCG
GCCCCCGGTTTCGTGGGGCCAGGGTCCAGCGGCTGCGCAGAGGCGGGGACCCCGGCCTCAT
GCACGGGAAGACTGTGTGATCACCGGGGCGAACAGCGGCCTGGGCCGCGCCACGGCCGCCG
AGCTACTGCGCCTGGGAGCGCGGGTGATCATGGGCTGCCGGGACCGCGCGCGCGCCGAGGAG
GCGGCGGGTCAGCTCCGCCGCGAGCTCCGCCAGGCCGCGGAGTGCGGCCAGAGCCTGGCGT
CAGCGGGGTGGGCGAGCTCATAGTCCGGGAGCTGGACCTCGCCTCGCTGCGCTCGGTGCGCG
CCTTCTGCCAGGAAATGCTCCAGGAAGAGCCTAGGCTGGATGTCTTGATCAATAACGCAGGG
ATCTTCCAGTGCCCTTACATGAAGACTGAAGATGGGTTTGAGATGCAGTTCGGAGTGAACCA
TCTGGGGCACTTTCTACTACCAATCTTCTCCTTGGACTCCTCAAAAGTTCACTCCCAGCA
GGATTGTGGTAGTTTCTTCCAACTTTATAAATACGGAGACATCAATTTTGATGACTTGAAC
AGTGAACAAAGCTATAATAAAAGCTTTTGTATTAGCCGGAGCAAAGTGGCTAACATTCTTTT
TACCAGGGAAGTAGCCCGCGCTTAGAAGGCACAAATGTACCGTCAATGTGTGTCATCCTG
GTATTGTACGGACAAATCTGGGGAGGCACATACATTCCACTGTTGGTCAAACCACTCTTC
AATTTGGTGTATGGGCTTTTTTCAAACTCCAGTAGAAGGTGCCAGACTTCATTATTT
GGCCTCTTCACCTGAGGTAGAAGGAGTGTGAGGAAGATACTTTGGGGATTGTAAAGAGGAAG
AACTGTTGCCCAAAGCTATGGATGAATCTGTTGCAAGAAAACCTCTGGGATATCAGTGAAGTG
ATGGTTGGCCTGCTAAAAAGGAACAAGGAGTAAAAGAGCTGTTTATAAACTGCATATCAG
TTATATCTGTGATCAGGAATGGTGTGGATTGAGAAGTTGTTACTTGAAGAAAAAGAATTTG
ATATTGGAATAGCCTGCTAAGAGGTACATGTGGGTATTTTGGAGTTACTGAAAAATTATTTT
TGGGATAAGAGAATTTCAAGCAAAGATGTTTTAAATATATATAGTAAGTATAATGAATAATAA
GTACAATGAAAAATACAATTATATTGTAATAATTATACTGGGCAAGCATGGATGACATATTA
ATATTGTGAGAAATTAAGTGACTCAAAGTGCTATCGAGAGGTTTTTCAAGTATCTTTGAGTT
TCATGGCCAAAGTGTTAACTAGTTTTTACTACAATGTTTGGTGTGTGTGGAAATTATCTGC
CTGGTGTGTGCACACAAGTCTTACTTGAATAAAATTACTGTTAC

FIGURE 122

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA58747

<subunit 1 of 1, 336 aa, 1 stop

<MW: 36865, pI: 9.15, NX(S/T): 2

MAVATAAAVLAALGGALWLAARRFVGPRVQRLRRGGDPGLMHGKTVLITGANSGLGRATAAE
LLRLGARVIMGCRDRARAEEAAGQLRRELQAAECGPEPGVSGVGELIVRELDLASLRSVRA
FCQEMLQEERLDVLINNAGIFQCPYMKTEDGFEMQFGVNLHGFLLTNLLLGLLKSSAPSR
IVVVSSKLYKYGDINFDDLNSEQSYNKSFCYSRSKLANILFTRELARRLEGTNVTNVNLHPG
IVRTNLGRHIHIPLLVKPLFNLVSWAFFKTPVEGAQTSIYLASSPEVEGVSGRYFGDCKEEE
LLPKAMDESVARKLWDISEVMVGLLK

Important features:

Signal peptide:

amino acids 1-21

Short-chain alcohol dehydrogenase family protein

amino acids 134-144, 44-56 and 239-248

N-glycosylation site.

amino acids 212-215 and 239-242

FIGURE 123

GGGGATTGTAAAGAGGAAGNACTGTGCCCAAAGNTATGGATGAATCTGTTGCAAGAAAATTN
TGGGATATCAGTGAAGTGATGGTTNGCCTGCTAAAAATAGGAACAAGGAGTAAAAGAGCTGTT
TATAAACTGCATATCAGTTATATCTGTGATCAGGAATGGTGTGGATTGAGAACTTGTTACT
TGAAGAAAAAGAATTTTGATATTGGAATAGCCTGNTAAGAGGNACATGTGGGTATTTTGGAG
TACTGAAAAATTATTTTGGGATAAGAGAATTCAGCAAAGATGTTTTAAATATATATAGT
AAGTATAATGAATAATAAGTACAATGAAAAATACAATTATATTGTAAAATTATAACTGGGCA
AGCATGGATGACATATTAATATTTGTCAGAATTAAGTGACTCAAAGTGCTATCGAGAGGTTT
TTCAAGTATCTTTGAGTTTCATGGCCAAAGTGTTAACTAGTTTTACTACAATGTTTGGTGTT
TGTGTGGAAATTATCTGCCTGGCTT

105701-105702

FIGURE 124

GAGAGGACGAGGTGCCGCTGCCTGGAGAATCCTCCGCTGCCGTCGGCTCCCGGAGCCCAGCC
CTTTCTTAACCAACCAACCTAGCCAGTCCCAGCCGACAGCGCTGTCCCTGTACGGAC
CCCAGCGTTACCAATGCATCCTGCCGCTCTTCTATCCTTACCCGACCTCAGATGCTCCCTTCT
GCTCCTGGTAACCTTGGGTTTTTACTCCTGTAACAACTGAAATAACAAGTCTTGCTACAGAGA
ATATAGATGAAATTTTAAACAATGCTGATGTTGCTTTAGTAAATTTTATGCTGACTGGTGT
CGTTTCAGTCAGATGTTGCATCCAATTTTGGAGGAAGCTCCGATGTCAATGAAGGAAGATT
TCCAAATGAAATCAAGTAGTGTGGCCAGAGTTGATTGTGATCAGCACTCTGACATAGCCC
AGAGATACAGGATAAGCAAATACCAACCCCTCAAATTGTTTCGTAATGGGATGATGATGAAG
AGAGAATACAGGGGTGAGCGATCAGTGAAAGCATTGGCAGATTACATCAGGCAACAAAAAG
TGACCCCATTCAGAAATTCGGGACTTAGCAGAAATCACCCTCTTGATCGCAGCAAAAGAA
ATATCATTGGATATTTTGAAGCAAAAGGACTCGGACAACCTATAGAGTTTTTGAACGAGTAGCG
AATATTTTGCATGATGACTGTGCCTTTCTTTCTGCATTGGGGATGTTTCAAAACCGGAAAG
ATATAGTGGCGACAACATAATCTACAAACCACCAGGCGATTCTGCTCCGGATATGGTGTACT
TGGGAGCTATGACAAATTTTGTGTGACTTACAATTGGATTCAAGATAAATGTGTTCTCTT
GTCCGAGAAATAACATTTGAAAATGGAGAGGAATTGACAGAAGAAGGACTGCCTTTTCTCAT
ACTCTTTCACATGAAAGAAGATACAGAAAGTTAGAAATATTCAGAAATGAAGTAGCTCGGC
AATTAATAAGTAAAAAGGTACAATAAACTTTTACATGCCGATTGTGACAAATTTAGACAT
CCTCTTCTGCACATACAGAAACTCCAGCAGATTGTCCTGTAATCGCTATTGACAGCTTTAG
GCATATGTATGTGTTTGGAGACTTCAAAGATGTATTAATTCCTGGAAACTCAAGCAATTCG
TATTTGACTTACATTTCTGGAAAACCTGCACAGAGAATTCCATCATGGACCTGACCCAACCTGAT
ACAGCCCCAGGAGAGCAAGCCCAAGATGTAGCAAGCAGTCCACCTGAGAGCTCCTTCAGAA
ACTAGCACCCAGTGAATATAGGTATACTCTATTGAGGGATCGAGATGAGCTTTTAAAAACTTG
AAAAACAGTTTGTAAGCCTTTCAACAGCAGCATCAACCTACGTGGTGAAATAGTAAACCTA
TATTTTCATAATTCTATGTGTATTTTATTTTGAATAAACAGAAAGAAATTTAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 125

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA57689

<subunit 1 of 1, 406 aa, 1 stop

<MW: 46927, pI: 5.21, NX(S/T): 0

MHPAVFLSLPDLRCSLLLLVTWVFTPVTTTEITSLATENIDEILNNADVALVNFYADWCRFSQ
MLHPIFEEASDVIKEEFPNENQVVFARVDCDQHS DIAQRYRISKYPTLKLFRNGMMMKREYR
GQRSVKALADYIRQQKSDPIQEIRDLAEITTLDRSKRNIIGYFEQKSDNYRVFERVANILH
DDCAFLSAFGDVSKPERYSGDNIIYKPPGHSAPDMVYLGAMTNFDVTYNWIQDKCVPLVREI
TFENGEELTEEGLPFLILFHMKEDTESLEIFQNEVARQLISEKGTINFLHADCDKFRHPLLH
IQKTPADCPVIAIDSFRHMYVFGDFKDVLI PGKLKQFVFDLHSGKLHREFHHGPDPTDTAPG
EQAQDVASSPPESSFQKLAPSEYRYTLLRDRDEL

Important features:

Signal peptide:

amino acids 1-29

Endoplasmic reticulum targeting sequence.

amino acids 403-406

Tyrosine kinase phosphorylation site.

amino acids 203-211

Thioredoxin family proteins

amino acids 50-66

FIGURE 126

ATTAAGGAAGAATTTCCAAATGAAAATCAAGTAGTNTTTGCCAGAGTNGATTGTGATCAGCA
CTCTGACATAGCCCAGAGATACAGGATAAGCAAATACCCAACCCTCAAATTGTTTCGTAATG
GGATGATGATGAAGAGAGAATACAGGGGTGAGCGATCAGTGAAAGCATTGGCAGATTA

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted October 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 127

AGAGGCCCTCTCTGGAAGTTGTCCCGGGTGTTCGCCGCNGGAGCCCGGGTCGAGAGGACNAGG
TGCCGCTGCCTGGAGAATCCTCCGCTGCCGTCGGCTCCCGGAGCCCAGCCCTTTCCTAAGCC
AACCCAACCTAGCCCNCTCCAGCCGCCAGCGCTGTCCCTGTCTCNGGANCACAGCGTNACC
ATGCATCCTGCCGTCTTCTATCCTTACCCGACCTCAGATGCTCCCTTCTGCTCCTGGTAAC
TTGGGTTTTTACTCCTGTAACAACGAAATAACNNGTCTTGATACNNAGAATATAGATGAAA
TTTTAAACNATGCTGATGTGGCTTTAGTCAATTTTATGCTGACTGGTGTCGTTTCAGTCAG
ATGTGGCATCCAATTTTTGAGGANGCTTCCGATGTCATTAAGGAAGAATTTCCAAATGAAAA
TCAAGTAGTGTGTCAGAGTTGATTGTGATCAGCACTCTGACATAGCCCAGAGATACAGGA
TAAGCAAATACCAACCCTCAAATTGTTTCGTAATGGGATGATGATGAAGAGAGAATACAGG
GGTCAGCGATCAGTGAAAGCATTGGCAGATTACATCAGGC

GCCACGCGTCCGATGGCGGTTACAGTTTCGCGGCCCTTCTGCTACATGCTGGCGCTGCTGCTCA
 CTGCCGCGCTCATCTTCTTCGCCATTGGCACATTATAGCATTGATGAGCTGAAGACTGAT
 TACAAGAATCCTATAGACCAGTGTAATACCTGAAATCCCCTTGTACTCCAGAGTACCTCAT
 CCACGCTTTTCTTCTGTGTATGTTTCTTTGTGCAGCAGAGTGGCTTACACTGGGTCTCAATA
 TGCCCTCTTGGCATATCATATTTGGAGGTATATGAGTAGACCAGTGTAGTGGCCAGGA
 CTCTATGACCTACAACCATCATGAATGCAGATATTCTAGCATATTGTGAGAAGGAAGGATG
 GTGCAAAATTAGCTTTTTATCTTCTAGCATTTTTTTACTACCTATATGGCATGATCTATGTTT
 TGGTGAGCTCTTAGAACAACACACAGAAGAATTGGTCCAGTTAAGTGCATGCAAAAAGCCAC
 CAAATGAAGGGATTCTATCCAGCAAGATCCTGTCCAAGAGTAGCCTGTGGAATCTGATCAGT
 TACTTTAAAAAATGACTCCTTATTTTTTAAATGTTTCCACATTTTGGCTGTGGAAGACTG
 TTTTCATATGTTATACTCAGATAAAGATTTTAAATGGTATTACGTATAAAATTAATATAAAAT
 GATTACCTCTGGTGTGACAGGTTTGAACTTGCACCTTCTTAAGGAACAGCCATAATCCTCTG
 AATGATGCATTAATTACTGACTGTCTAGTACATTGGAAGCTTTTGTTTATAGGAACCTTGTA
 GGGCTCATTTTGGTTTCATTGAAACAGTATCTAATTATAAATTAGCTGTAGATATCAGGTGC
 TTCTGATGAAGTGAAAAATGTATATCTGACTAGTGGGAACTTCATGGGTTTCTCATCTGTC
 ATGTCGATGATTATATATGGATACATTTACAAAAATAAAAAGCGGGAATTTTCCCTTCGCTT
 GAATATTATCCCTGTATATTGCATGAATGAGAGATTTCCCATATTTCCATCAGAGTAATAAA
 TATACTTGCTTTAATTCTTAAGCATAAGTAACATGATATAAAAAATATATGCTGAATTACTT
 GTGAAGAATGCATTTAAAGCTATTTTAAATGTGTTTTTATTGTGAAGACATTACTTATTAAG
 AAATTGGTTATTATGCTTACTGTTCTAATCTGGTGGTAAAGGTATTTCTAAGAATTTGCAGG
 TACTACAGATTTTCAAACCTGAATGAGAGAAAAATGTATAACCATCCTGCTGTTCCTTTAGT
 GCAATACAATAAACTCTGAAATTAAGACTC

FIGURE 129

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA23330

<subunit 1 of 1, 144 aa, 1 stop

<MW: 16699, pI: 5.60, NX(S/T): 0

MAFTFAAFCYMLALLLTAAIIFFAIWHIIAFDELKTDYKNPIDQCNTLNPLVLPEYLIHAFF
CVMFLCAAEWLTGLNMPLLAYHIWRYMSRPVMSGPGLYDPTTIMNADILAYCQKEGWCKLA
FYLLAFFYYLYGMIYVLVSS

Important features:

Signal peptide:

amino acids 1-20

Type II transmembrane domain:

amino acids 11-31

Other transmembrane domain:

amino acids 57-77 and 123-143

2025-10-10 10:10:10

ATTATAGCATTGTGATGAGCTGAAGACTGATTACAAGATCCTATAGACCAGTGTAAATACCCGT
AATCCCCCTTGTACTCCAGAGTACCTCATCCACGCTTTCTTCTGTGTCATGTTTCTTTGTGC
AGCAGAGTGGCTTACACTGGGTCTCAATATGCCCTCTTGGCATAATCATATTTGGAGGTATA
TGAGTAGACCAGTGTAGAGTGGCCCAGGACTCTATGACCTTACAACCATCATGAATGCAGAT
ATTCTAGCATATTGTGAGAAGGAAGGATGGTGCAAATTAGCTTTTTATCTTCTAGCATTTTT
TTACTACCTATATGGCATGATCTATGTTTTGGTGAGCTCTTAGAACACACAGAAGAATT
GGTCCAGTTAAGTGCATGCAAAAAGCCACCAATGAAGGGATTCTATCCAGCAAGATCCTGT
CCAAGAGTAGCCTGTGGAATCTGATCAGTTACTTTAAAAAATG

FIGURE 131

CGGACGCGTGGGGGAAACCTTCCGAGAAAACAGCAACAAGCTGAGCTGCTGTGACAGAGGG
GAA CAAGATGCGCGCGCCGAAGGGGAGCCTCTGGGTGAGGACCCAACCTGGGGCTCCCGCCGC
TGCTGCTGCTGACCATGGCCTTGGCCGGAGGTTCTGGGGACCGCTTCGGCTGAAGCATTGTGAC
TCGGTCTTGGGTGATACGGCGCTTGTGCCACCGGCCTGTCTAGTTGACCTACCCCTTGCACAC
CTACCCTAAGGAAGAGGAGTTGTACGCATGTCAGAGAGGTTGCAGGCTGTTTTCAATTTGTC
AGTTTGTGGATGATGGAATTGACTTAAATCGAACTAAATTGGAATGTGAATCTGCATGTACA
GAAGCATATTTCCCAATCTGATGAGCAATATGCTTGCCATCTTGGTTGCCAGAATCAGCTGCC
ATTGCTGAACTGAGACAAGAACAATTATGTCCCTGATGCCAAAATGCACCTACTCTTTC
CTCTAACTCTGGTGAGGTCATTCTGGAGTGACATGATGGACTCCGCACAGAGCTTCATAACC
TCTTCATGGACTTTTTATCTTCAAGCCGATGACGGAATAAGTTATATTCCAGTCTAAGCC
AGAAATCCAGTACGCACCACATTGGAGCAGGAGCCTACAAATTTGAGAGAATCATCTCTAA
GCAAAATGTCTTATCTGCAAATGAGAAATTCACAAGCGCACAGGAATTTTCTGAAGATGGA
GAAAGTGATGGCTTTTTAAGATGCCTCTCTCTTAACTCTGGGTGGATTTTAACTACAACCTCT
TGTCCTCTCGGTGATGGTATTGCTTTGGATTTGTTGTGCAACTGTTGCTACAGCTGTGGAGC
AGTATGTTCCCTCTGAGAAGCTGAGTATCTATGGTGACTTGGAGTTTATGAATGAACAAAAG
CTAAACAGATATCCAGCTTCTTCTCTTGTGTTGTTAGATCTAAACTGAAGATCATGAAGA
AGCAGGGCCTCTACCTACAAAAGTGAATCTTGCTCATTCTGAAATTTAAGCATTTTTCTTTT
AAAAGACAAGTGTAATAGACATCTAAAATTCACCTCCTCATAGAGCTTTTAAAATGGTTTCA
TTGGATATAGGCCTTAAGAAATCACTATAAAATGCAAAATAAGTTACTCAAATCTGTG

FIGURE 132

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA26847

<subunit 1 of 1, 323 aa, 1 stop

<MW: 36223, pI: 5.06, NX(S/T): 1

MAAPKGSILWVRTQLGLPPLLLLLTALAGSGTASAEAFDSVLGDTASCHRAQQLTYPLHTYP
KEEELYACQRGCRFLFSICQFVDDGIDLNRKLECESECTEAYSQSDEQYACHLGCONQLPFA
ELRQEQQLMSLMPKMHLFLPLTLVRSFWSMMDSAQSFITSSWTFYLQADDGKIVIFQSKPEI
QYAPHLEQEPTNLRESSLSKMSYLMRNSQAHRNFLEDGESDGLRCLSLNSGWILTTTLVL
SVMVLLWICCATVATAVEQYVPSEKLSIYGDLEFMNEQKLNRYPASSLVVVRSKTEDHEEAG
PLPTKVNLAHSEI

Important features:

Signal peptide:

amino acids 1-31

Transmembrane domain:

amino acids 241-260

N-glycosylation site.

amino acids 90-93

FIGURE 133

TTGGGTGATACGGCGTCTTGCCACCGGGCCTGTCAGTTGACCTACCCCTTGCAACCTACCC
TAAGGAAGAGGAGTTGTACGCATGTCAGAGAGGTTGCAGGCTGTTTTCAATTTGTCAGTTTG
TGGATGATGGAATTGACTTAAATCGAACTAAATTGGAATGTGAATCTGCATGTACAGAAGCA
TATTCCCAATCTGATGAGCAATATGCTTGCCATCTTGGTTGCCAGAATCAGCTGCCATTGCG
TGAAC TGAGACAAGAACAAC TTTATGTCCCTGATGCCAAAAATGCACCTACTCTTTCCTCTAA
CTCTGGTGAGGTCATTCTGGAGTGACATGATGGACTCCGC

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1489
1490
1491
1492
1493
1494
1495
1496
1497
1498
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1598
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697
1698
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1760
1761
1762
1763
1764
1765
1766
1767
1768
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1789
1790
1791
1792
1793
1794
1795
1796
1797
1798
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2040
2041
2042
2043
2044
2045
2046
2047
2048
2049
2050
2051
2052
2053
2054
2055
2056
2057
2058
2059
2060
2061
2062
2063
2064
2065
2066
2067
2068
2069
2070
2071
2072
2073
2074
2075
2076
2077
2078
2079
2080
2081
2082
2083
2084
2085
2086
2087
2088
2089
2090
2091
2092
2093
2094
2095
2096
2097
2098
2099
2100
2101
2102
2103
2104
2105
2106
2107
2108
2109
2110
2111
2112
2113
2114
2115
2116
2117
2118
2119
2120
2121
2122
2123
2124
2125
2126
2127
2128
2129
2130
2131
2132
2133
2134
2135
2136
2137
2138
2139
2140
2141
2142
2143
2144
2145
2146
2147
2148
2149
2150
2151
2152
2153
2154
2155
2156
2157
2158
2159
2160
2161
2162
2163
2164
2165
2166
2167
2168
2169
2170
2171
2172
2173
2174
2175
2176
2177
2178
2179
2180
2181
2182
2183
2184
2185
2186
2187
2188
2189
2190
2191
2192
2193
2194
2195
2196
2197
2198
2199

FIGURE 134

CACACTGGCCGGATCTTTAGAGTCCTTTGACCTTGACCAAGGGTCNGGAAAACAGCAACAA
GCTGAGCTGCTGTGACAGAGGGAACAAGATGGCGGCGCCGAAGGGAGCCTTTGGGTGAGGAC
CCAACTGGGGCTCCCGCGCTGCTGCTGCTGACCATGGCCTTGCCCGGAGGTTCGGGGACCG
CTTCGGCTGAAGCATTGTGACTCGGTCTTGGGTGATACGGCGTCTTGCCACCGGGCCTGTCAG
TTGACCTACCCCTTGACACCTACCCCTAAGGAAGAGGAGTTGTACGCATGTCAGAGAGGTTG
CAGGCTGTTTTCAATTTGTCACTTTGTGGATGATGGAATTGACTTAAATCGAACTAAATTGG
AATGTGAATCTGCATGTACAGAAGCATATTCCAATCTGATGAGCAATATGCTTGCCATCTT
GGTTGCCAGAATCAGCTGCCATTGCTGAACTGAGACAAGAACAACCTTATGTCCCTGATGCC
AAAAATGCACCTACTCTTTCCTCTAACTCTGGTGAGGTCATTCTGGAGTGACATGATGGACT
CCGC

FIGURE 135

GCGAGGTGGCGATCGCTGAGAGGCAGGAGGGCCGAGGCGGGCCTGGGAGGCGGCCCGGAGGT
GGGGCGCCGCTGGGGCCGGCCCGCACGGGCTTCATCTGAGGGCGCACGGCCCGCGACCGAGC
GTGCGGACTGGCCCTCCCAAGCGTGGGGCGACAAGCTGCCGGAGCTGCAATGGGCGCGGCTG
GGGATTCTTGTTTGGCCCTCTGGGCGCCGTGTGGCTGCTCAGCTCGGGCCACGGAGAGGAGC
AGCCCCCGGAGACAGCGGCACAGAGGTGCTTCTGCCAGGTTAGTGGTTACTTGGATGATTGT
ACCTGTGATGTTGAAACCATTGATAGATTTAATAACTACAGGCTTTTCCCAAGACTACAAAA
ACTTCTTGAAAGTGACTACTTTAGGTATTACAAGGTAAACCTGAAGAGGCCGTGTCTTTCT
GGAATGACATCAGCCAGTGTGGAAGAAGGACTGTGCTGTCAAACCATGTCAATCTGATGAA
GTTCTGTATGGAATTAATCTGCGAGCTACAAGTATTCTGAAGAAGCCAATAATCTCATTGA
AGAATGTGAACAAGCTGAACGACTTGGAGCAGTGGATGAATCTCTGAGTGAGGAAACACAGA
AGGCTGTTCTTCACTGGACCAAGCATGATGATTCTTCAGATAACTTCTGTGAAGCTGATGAC
ATTCAGTCCCCTGAAGCTGAATATGTAGATTGCTTCTTAATCCTGAGCGCTACACTGGTTA
CAAGGGACCAGATGCTTGAAAAATATGGAATGTCTACGAAGAAAACCTGTTTTAAGCCAC
AGACAATTAAAAGACCTTTAAATCCTTTGGCTTCTGGTCAAGGGACAAGTGAAGAGAACT
TTTTACAGTTGGCTAGAAGTCTCTGTGTAGAAAAAGAGCATTCTACAGACTTATATCTGG
CCTACATGCAAGCATTAAATGTGCATTGAGTGCAAGATATCTTTTACAAGAGACCTGGTTAG
AAAAGAAATGGGGACACAACATTACAGAATTTCAACAGCGATTGATGGAATTTTGACTGAA
GGAGAAGGTCCAAGAAGGCTTAAGAACTTGATTTTCTCTACTTAATAGAACTAAGGCTTT
ATCCAAAGTGTTACCATTCTTCGAGCGCCAGATTTTCAACTCTTTACTGGAATAAAATTC
AGGATGAGGAAAAACAAATGTTACTTCTGGAAATACTTCATGAAATCAAGTCATTTCTTTG
CATTTTGATGAGAATTCATTTTTGTCTGGGGATAAAAAAGACACAAACTAAAGGAGGA
CTTTCGACTGCATTTTAGAAATATTTCAAGAATTATGGATTGTGTTGGTTGTTTTAAATGTC
GTCGTGGGGAAAGCTTCAGACTCAGGGTTTGGGCACTGCTCTGAAGATCTTATTTTCTGAG
AAATTGATAGCAAATATGCCAGAAAGTGGACCTAGTTATGAATTCATCTAACCAGACAAGA
AATAGTATCATTATTCACGCAATTTGGAAGAATTTCTACAAGTGTGAAAGAATTAGAAAACT
TCAGGAACCTGTTACAGAATATTCATTAAAGAAAAACAGCTGATATGTGCCTGTTTCTGGAC
AATGGAGGCGAAAGAGTGGAATTTCAATCAAGGCATAATAGCAATGACAGTCTTAAGCCAA
ACATTTTATATAAAGTTGCTTTTGTAAAGGAGAATTATATGTTTAAAGTAAACACATTTT
AAAAATTGTGTTAAGTCTATGTATAATACTACTGTGAGTAAAGTAATACTTTAATAATGTG
GTACAAATTTTAAAGTTAATATTGAATAAAAGGAGGATTATCAAATTAATAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAA

CITIZENSHIP

<MW: 54393, pI: 5.63, NX(S/T): 2

MGRGWGFLFGLLAGVWLLSSSGHGEEQPPETAARCFQVSGYLDCTCDVETIDRFNNYRLF
PRLQKLLSDFRYRYKVNLRKPCPFWNDISQCGRRDCAVKPCQSDVPDGIKSASYKYSEEA
NNLIEECEQAERLGAVDESLSSETQKAVLQWTKHDDSSDNFCEADDIQSPEAYVDLLNPE
RYTGYKGPDAWKIWNVIYEENCFKPQTIKRPLNPLASGQGTSEENTFYSWLEGLCVEKRAFY
RLISGLHASINVHLSARYLLQETWLEKKWGHNITFQQRFDGILTEGEGPRRLKNLYFLYLI
ELRALSKVLPFFERPDFQLFTGNKIQDEENKMLLLEILHEIKSFPLHFDENSFFAGDKKEAH
KLKEDFRLHFRNISRIMDCVGCFCRLWGKLQTQGLGTALKILFSEKLIANMPESGPSYEFH
LTROEIVSLFNAFGRISTSVKELENFRNLLONIH

Important features:

Signal peptide:

amino acids 1-23

N-glycosylation site.

amino acids 280-283 and 384-387

Amidation site.

amino acids 94-97

Glycosaminoglycan attachment site.

amino acids 20-23 and 223-226

Important features:

Signal peptide:

amino acids 1-23

N-glycosylation site.

amino acids 280-283 and 384-387

Amidation site.

amino acids 94-97

Glycosaminoglycan attachment site.

amino acids 20-23 and 223-226

Aminotransferases class-V pyridoxal-phosphate

amino acids 216-222

Interleukin-7 proteins

amino acids 338-343

FIGURE 137

GCTGGAATATGGATGTCATCTACGAGAACTGTTTTAAGCCACAGACAATTAAAAGACCTT
TAAATCCTTTGGCTTCTGGTCAAGGGACAAGTGAAGAGNACACTTTTTACAGTTGGCTAGAA
GGTCTCTGTGTAGAAAAAGAGCATTCTACAGACTTATATCTGGCCTACATGCAAGCATTA
TGTGCATTTGAGTGCAAGATATCTTTTACAAGAGACCTGGTTAGAAAAGAAATGGGGACACA
ACATTACAGAATTTNAACAGCGATTTGATGGAATTTTGACTGAAGGAGAAGGTCCAAGAAGG
CTTAAGAACTTGTATTTTCTCTACTTAATAGAACTAAGGGCTTTATCCAAAGTGTACCATT
CTTNGAGCGCCAGATTTTCACTNTTTACTGGAAATAAAATTCAGGATGAGGNAACAAAA
TGTTACTTTTGAAATACTTCATGAAATCAAGTCATTTCTTTGCATTTTGATGAGAATTCA
TTTTTTTGCTG

FIGURE 138

CGGACGCGTGGGCGGACGCGTGGGCGGACGCGTGGGTGGGAGGGGGCAGGATGGGAGGGAA
AGTGAAGAAAACAGAAAAGGAGAGGGACAGAGGCCAGAGGACTTCTCATACTGGACAGAAAC
CGATCAGGCATGGAACTCCCCCTCGTCACTCACCTGTTCTTGCCCTGGTGTTCCTGACAGG
TCTCTGCTCCCCCTTTAACTGGATGAACATCACCCACGCTATTCCCAGGGCCACCAGAAG
CTGAATTTGGATACAGTGTCTTACAAATGTTGGGGGTGGACAGCGATGGATGCTGGTGGGC
GCCCCCTGGGATGGGCCTTCAGGCGACCGAGGGGGGACGTTTATCGCTGCCCTGTAGGGGG
GGCCCAATGCCCCATGTGCCAAGGGCCACTTAGGTGACTACCAACTGGGAAATTCATCTC
ATCCTGCTGTGAATATGCACCTGGGGATGTCTCTGTTAGAGACAGATGGTGTGGGGGATT
ATGGTGAGCTAAGGAGAGGGTGGTGGCAGTGTCTCTGAAGGTCCATAAAAGAAAAAGAGAA
GTGTGGTAAGGGAAAATGGTCTGTGTGGAGGGTCAAGGAGTTAAAAACCTAGAAAGCAA
AGGTAGGTAATGTCAGGGAGTAGTCTTTCATGCCTCCTTCAACTGGGAGCATGTTCTGAGGGT
GCCCTCCCAAGCCTGGGAGTAACTATTTCCCCCATCCCCAGGCCTGTGCCCTCTCTGGTCT
CGTGCTTGTGGCAGCTCTGTCTTCAGTTCTGGGATATGTGCCCGTGTGGATGCTTCATTCCA
GCCTCAGGGAAGCCTGGCACCCTGCCCCAACGTGAGCCAGAGGAAGGCTGAGTACTTGGTT
CCCAGAAGGAGATACTGGGTGGGAAAAAGATGGGGCAAAGCGGTATGATGCCTGGCAAAGG
CCTGCATGGCTATCCTCATTTGTACCTAATGTGCTTGCAAAGCTCCATGTTTCCTAACAGA
TTCAGACTCCTGGCCAGGTGTGGTGGCCACACCTGTAATTCTAGCACTTTGGGAGGCCAAG
GTGGGCAGATCACTTGAGGTGAGGAGTTCAAGACCAGCCTGGCCAACATGGTGAAACTCCAT
CTCTACTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA
ATCTACTCGGAGGCTAAGACAGGAGACTCTCACTTCAACCAGGAGGTGGAGGTGCGGTG
AGCCAAGATTGTGCTCTGCACTCTAGCGTGGGTGACAGAGTAAGCGAGACTCCATCTCAA
AATAATAATAATAATAATTAGACTCCTTATCAGGAGTCCATGATCTGGCCTGGCACAGTAA
CTCATGCCGTGAATCCCAACATTTTGGGAGGCCAACGAGGAGGATTGCTTGAGGTCTGGAG
GTTTGAGACCAGCCTGGGCAACATAGAAAGACCCCATCTCTAAATAAATGTTTTAAAAAT

FIGURE 139

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA57039

><subunit 1 of 1, 124 aa, 1 stop

><MW: 13352, pI: 5.99, NX(S/T): 1

MELPFVTHLFLPLVFLTGLCSPPNLDEHHPRLFPGPPEAEFGYSVLQHVGGGQRWMLVGAPW

DGPSGDRRGDVYRCPVGGAHNAPCAKGHLGDYQLGNSSHPAVNMHLGMSLLETGDGGGFMVS

Important features:

Signal peptide:

amino acids 1-22

Cell attachment sequence.

amino acids 70-73

N-glycosylation site.

amino acids 98-101

Integrins alpha chain proteins

amino acids 67-81

FIGURE 140

CACAGTTCCCCACCATCACTCNTCCCATTCCCTTCCAACTTTATTTTGTAGCTTGCCATTGGGA
GGGGGCAGGATGGGAGGGAAAGTGAAGAAAACAGAAAAGGAGAGGGACAGAGGCCAGAGGAC
TTCTCATACTGGACAGAAACCGATCAGGCATGGAACTCCCCTTCGTCACTCACCTGTTCTTG
CCCCTGGTGTTCCCTGACAGGTCTCTGCTCCCCCTTTAACCTGGATGAACATCACCCACGCCT
ATTCCCAGGGCCACCAGAAGCTGAATTTGGATACAGTGTCTTACAACATGTTGGGGGTGGAC
AGCGATGGATGCTGGTGGGCGCCCCCTGGGATGGGCCTTCAGGCGACCGGAGGGGGGACGTT
TATCGCTGCCCTGTAGGGGGGGCCACAATGCCCCATGTGCCAAGGGCCACTTAGGTGACTA
CCAAC TGGGAAATTCATCTCATCTGCTGTGAATATGCACCTGGGGATGTCTCTGTTAGAGA
CAGATGGTGATGG

FIGURE 141

AAAGTTACATTTTCTCTGGAACCTCTCCTAGGCCACTCCCTGCTGATGCAACATCTGGGTTTG
GGCAGAAAGGAGGGTGCTTCGGAGCCCGCCCTTTCTGAGCTTCTGGGCCGGCTCTAGAACA
ATTCAAGCTTCGCTGCGACTCAGACCTCAGCTCCAACATATGCATTCTGAAGAAAGATGGCT
GAGATGGACAGAATGCTTTATTTTGGAAAGAAACAATGTTCTAGGTCAAACCTGAGTCTACCA
AATGCAGACTTTCACAATGGTTCTAGAAGAAATCTGGACAAGTCTTTTCATGTGGTTTTTCT
ACGCATTGATTCATGTTTGCTCAGATGAAGTGGCCATTCTGCTGCCCTCAGAACCTC
TCTGTACTCTCAACCAACATGAAGCATCTCTTGATGTGGAGCCAGTGATCGCGCTTGAGA
AACAGTGTACTATTCTGTCGAATACCAGGGGGAGTACGAGAGCCTGTACACGAGCCACATCT
GGATCCCCAGCAGCTGGTGCTCACTCACTGAAGTCCTGAGTGTGATCTCACTGATGACATC
ACGGCCACTGTGCCATACAACCTTCGTGTGAGGGCCACATTGGGCTCAGACCTCAGCCTG
GAGCATCCTGAAGCATCCCTTTAATAGAACTCAACCATCCTTACCCGACCTGGGATGGAGA
TCACCAAGATGGCTTCCACCTGGTTATTGAGCTGGAGGACCTGGGGCCCCAGTTTGAGTTC
CTTGTTGGCTTACTGGAGGAGGGAGCCTGGTGCCGAGGAACATGTCAAATGGTGAGGATGG
GGGTATTCAGTGCACCTAGAAACCATGGAGCCAGGGGCTGCATACTGTGTGAAGGCCCAGA
CATTCTGTAAGGCCATTGGGAGGTACAGCGCTTCAGCCAGACAGAATGTGTGGAGGTGCAA
GGAGAGGCCATTCCCTGGTACTGGCCCTGTTTGCCCTTGTGTGGCTTCATGCTGATCCTTGT
GGTCGTGCCACTGTTTCGTCTGAAAAATGGGCCGGCTGCTCCAGTACTCCTGTTGCCCGTGG
TGGTCCTCCAGACACCTTGAAATAACCAATTCACCCAGAAGTTAATCAGTGCAGAAAG
GAGGAGGTGGATGCCGTGTGCCACGGCTGTGATGTCTCCTGAGGAACTCCTCAGGGCCTGGAT
CTCATAGGTTTGCGGAAGGGCCAGGTGAAGCCGAGAACCTGGTCTGCATGACATGGAAACC
ATGAGGGGACAAGTTGTGTTTCTGTTTTCGCCACCGGACAAGGGATGAGAGAAGTAGGAAGA
GCCTGTTGTCTACAAGTCTAGAAGCAACCATCAGAGGCAGGGTGTTTGTCTAACAGAACAC
TGACTGAGGCTTAGGGGATGTGACCTCTAGACTGGGGCTGCCACTTGCTGGCTGAGCAACC
CTGGGAAAAGTGACTTCATCCCTTCGGTCCTAAGTTTTCTCATCTGTAATGGGGGAATTACC
TACACACCTGCTAAACACACACACAGAGTCTCTCTATATATACACACGTACACATAAA
TACACCCAGCACTTGCAAGGCTAGAGGGAACTGGTGACACTCTACAGTCTGACTGATTGAG
TGTTTTCTGGAGAGCAGGACATAAATGTATGATGAGAAATGATCAAGGACTCTACACACTGGGT
GGCTTGAGAGGCCACTTTCAGAAATAATCCTTGAGAGAAAAGGAATCATGGGAGCAATGG
TGTTGAGTTCACTTCAAGCCCAATGCCGTGCAGAGGGGAATGGCTTAGCGAGCTCTACAGT
AGGTGACCTGGAGGAAGGTACAGCCACACTGAAATGGGATGTGCATGAACACGGAGGATC
CATGAACACTGTAAAGTGTGACAGTGTGTGCACACTGCAGACAGCAGGTGAAATGTATGT
GTGCAATGCGACGAGAATGCAGAAGTCAGTAACATGTGCATGTTTGTGTGCTCCTTTTTC
TGTTGGTAAAGTACAGAATTCAGCAAATAAAAAGGGCCACCCTGGCCAAAAGCGGTAAAAAA
AAAAA

037543-1450

FIGURE 142

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA57033

<subunit 1 of 1, 311 aa, 1 stop

<MW: 35076, pI: 5.04, NX(S/T): 2

MQTFTMVLEEIIWTSLFMWFFYALIPCLLTDEVAILPAPQNLSVLSTNMKHLMLWSPVIAPGE
TVYYSVEYQGEYESLYTSHIWIPSSWCSLTEGPECDVTTDITATVPYNLVRATLGSQTSAN
SILKHPFNRNSTILTRPGMEITKDGPHLVIELEDLGPQFEFLVAYWRREP GABEHVKMVRSG
GIPVHLETMEPGAAYCVKAQTFVKAIGRYSAFSQTECVEVQGEAIPVLALFAFVGFMILILV
VVPLFVWKMGRLLQYSCCPVVVLPDTLKITNSPQKLISCRREEVDACATAVMSPHEELRAWIS

Important features:

Signal peptide:

amino acids 1-29

Transmembrane domain:

amino acids 230-255

N-glycosylation site.

amino acids 40-43 and 134-137

Tissue factor proteins.

amino acids 92-119

Integrins alpha chain proteins

amino acids 232-262

Figure 1

TCTCTGCTGATGCACATCTGGGTTTGGCAAAAGGAGGTTGCTTCGAGCCGCCCCTTTCTAGCTT
CCTGGCCGGCTCTAGAACAAATTGAGGCTTCGCTCGCAGTACAGCTCAGCTCCAACATATGCA
TTCTGAAGAAAGATGGCTGAGATGACAGAATGCTTTATTTTGGAAAGAAACAATGTTCTAGG
TCAAACCTGAGTCTACCAAAATGCAGACTTTTACAATGGTTCTAGAAGAAATCTGGACAAGTCT
TTTTCATGTGGTTTTTCTACGCATTGATTCCATGTTTGGCTCACAGATGAAGTGCCATTCTGC
CTGCCCTCAGAACCTCTCTGTACTCTCAACCAACATGAAGCATCTCTTGTATGTGGAGCCCA
GTGATCGCGCTGGAGAAACAGTGTACTATTCTGTCGAATACCAGGGGGAGTACGAGAGCCT
GTACACGAGCCACATCTGGATCCCCAGCAGCTGGTGCTCACTCACTGAAGGTCCTGAGTGTG
ATGTCACTGATGACATCACGGCCACTGTGCCATACAACCTTTGTGTACAGGGCCACATTGGGC
TCACAGACCTCAGCCTGGAGCATCTGAAGCATCCCTTTAATAGAAACTCAACCATCTTTAC
CCGACCTGGGATGGAGATCACCAAAGATGGCTTNCACCTGGTTATTGAGCTGGAGGACCTGG
GGCCCCAGTTTGAGTTTCCTTGTGGCCTANTGGAGGAGGGGGCGAACCCCTTGCGGCGCAAGGG
GTTNCGCAACCCCTTGCGGCCGCTGGGGTATCTCTCGAGAAAAGAGAGGCCCAATATGACCC
ACATACTCAATATGGACGAANTGCTATTGTCCACCTGTTTGAGTGGCGCTGGGTTGAT

CCACGCGTCCGCCACGCGTCCGAGGGACAAGAGAGAAGAGAGACTGAAACAGGGAGAAGA
GGCAGGAGAGGAGGAGGTGGGGAGAGCACGAAGCTGGAGGCCGACACTGAGGGAGGGCGGGA
GGAGGTGAAGAAGGAGAGAGGGGAGAAGAGGCAGGAGCTGGAAGGAGAGAGGGAGGAGGAG
GAGGAGATGCGGGATGAGACCTGGAGTTAGGTGGCTTGGGAGAGCTTAATGAAAAGAGAAC
GGAGAGGAGGTGTGGGTTAGGAACCAAGAGGTAGCCCTGTGGGCAGCAGAAGGCTGAGAGGA
GTAGGAAGATCAGGAGCTAGAGGGAGACTGGAGGGTTCCGGGAAAAGAGCAGAGGAAAGAGG
AAAGACACAGAGAGACGGGAGAGAGAAGAAGAGTGGGTTTGAAGGGCGGATCTCAGTCCCTG
GCTGCTTTGGCATTGTGGGAACTGGGACTCCCTGTGGGGAGGAGAGGAAAGCTGGAAGTCCT
GGAGGACAGGGTCCCAGAAAGAGGGGACAGAGGAGCTGAGAGAGGGGGCAGGGCGTTGGG
CAGGGGTCCCTCGGAGGCCCTCTGGGGATGGGGCTGCAGCTCGTCTGAGCGCCCTCGAGC
GCTGGTACTCTGGGCTGCACTGGGGGCAGCAGCTCACATCGGACCAGCACCTGACCCCGAGG
ACTGTTGGAGCTACAAGGATAATCTCCAGGGAAACTTCGTGCCAGGGCCTCCTTTCTGGGGC
CTGGTGAATGCAGCGTGGAGTCTGTGTGCTGTGGGGAAGCGGCAGAGCCCCGTGGATGTGGA
GCTGAAGAGGGTTCTTTATGACCCCTTTCTGCCCCATTAAAGGCTCAGCACTGGAGGAGAGA
AGCTCCGGGGAACTTGTACAAACACCGGCCGACATGTCTCCTTCTGCTGCACCCCGACCT
GTGGTCAATGTGTCTGGAGGTCCCTCCTTTACAGCCACCGACTCAGTGAACTCGGGCTGCT
GTTTGGAGCTCGCGACGGAGCCGGCTCGGAACATCAGATCAACCACAGGGCTTCTCTGCTG
AGGTGCAGCTCATTCACTTCAACCAGGAACTCTACGGGAATTTAGCGCTGCCTCCCGCGGC
CCCAATGGCCTGGCCATTCTCAGCCTCTTTGTCAACGTTGCCAGTACCTTAACCCATTCTCT
CAGTCGCCTCCTTAACCGCGACACCATCACTCGCATCTCCTACAAGAATGATGCTACTTTC
TTCAAGACCTGAGCCTGGAGCTCCTGTTCCCTGAATCCTTCGGCTTCATCACCTATCAGGGC
TCTCTCAGCACCCCGCCCTGCTCCGAGACTGTCACTGGATCTCATTGACCGGGCCCTCAA
TATCACCTCCCTTCAGATGCACTCCCTGAGACTCCTGAGCCAGAATCTCCATCTCAGATCT
TCCAGAGCCTCAGCGGTAACAGCCGGCCCCCTGCAGCCCTTGGCCACAGGGCACTGAGGGGC
AACAGGGACCCCGGCACCCCGAGAGGCGCTGCCGAGGCCCAACTACCGCCTGCATGTGGA
TGGTGTCCCCCATGGTCGCTGAGACTCCCTTCGAGGATTGCACCCGCCCGTCTTAAGCCTC
CCCACAAGGCGAGGGGAGTTACCCTTAAAAAAGGCTATTAAAGGGACAGAATACTTA

FIGURE 145

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA34353

<subunit 1 of 1, 328 aa, 1 stop

<MW: 36238, pI: 9.90, NX(S/T): 3

MGAAARLSAPRALVLWAALGAAAHIGPAPDPEDWWSYKDNLQGNFVPGPPFWGLVNAAWSLC
AVGKRQSPVDVELKRVLYDPFLPPLRLSTGGEKLRGTLTYNTRHVSFLPAPRPVVNVSGGPL
LYSHRLSELRLFLGARDGAGSEHQINHQGFSAEVQLIHFNQELYGNFSAASRGPNGLAILSL
FVNVASTSNPFLSRLLRDITITRISYKNDAFLQDLSLELLFPESFGFITYQGSLSLTPPCSE
TVTWILIDRALNITSLQMHSRLRLSQNPSSQIFQSLSGNSRPLQPLAHRALRGNRDPRHPR
RCRGPNYRLHVDGVPHGR

Important features:

Signal peptide:

amino acids 1-23

Transmembrane domain:

amino acids 177-199

N-glycosylation site.

amino acids 118-121, 170-173 and 260-263

Eukaryotic-type carbonic anhydrases proteins

amino acids 222-270, 128-164 and 45-92

FIGURE 146

GGCGCCTGGTTCTGCGCGTACTGGCTGTACGGAGCAGGAGCAAGAGTTCGCGCCGACGCTCCGCGCCGAGCCTC
 GTTCGTGTCCCCGCCCTCGCTCCTGCAGCTACTGCTCAGAAACGCTGGGGCGCCACCTTGGCAGACTAACGAA
 GCAGCTCCCTTCCACCCCAACTGCAGGTCTAATTTTGGACGCTTTGCTTCCAGTTTGGAGGAGC
 CGCAGAGCGGAGGCTCGCGTATTTCTGAGTCAGCACCCACGTCGCCCCGAGCGCTCGGTGCTCAGGCCCTTC
 GCGAGCGGGGCTCTCCGTCTGCGGTCCTTGTGAAGGCTCTGGGCGGCTGCAGAGGCGCGCGCTCCGGTTTGGCT
 CACCTCTCCAGGAACTTCACACTGGAGAGCCAAAAGGAGTGAAGAGCCTGTCTTGGAGATTTTCTTGGGGAA
 ATCCTGAGGTCAITTCATTATGAAGTGTACGCGCGGGAGTGGCTCAGAGTAACACAGTGTCTGTTTATGGCTAGA
 GCAATTCAGGCCATGGTGGTTCCCAATGCCACTTTATTGGAGAACTTTTGGAAAAATACATGGATGAGGATGGT
 GAGTGGTGGATAGCCAAACACGAGGGGAAAAGGGCCATCACAGACAATGACATCAGAGTATTTTGGACCTTCAT
 AATAAATTACGAAGTCAGGTGTATCTCAACAGCCTCTAATATGGAGTATATGACATGGGATGTAGAGCTGGAAAGA
 TCTGCAGAATCCTGGGCTGAAAGTGTCTTGTGGGAACATGGACCTGCAAGCTTGTCTCCATCAATTGGACAGAAT
 TTGGGAGCACACTGGGGAAGATATAGGCCCCCGACGCTTTCATGTACAATCGTGGTATGATGAAGTGAAGACTTT
 AGCTACCCATATGAACATGAATGCAACCCATATTGTCCATTGAGTGTCTTGGCCCTGTATGTACACATTATACA
 CAGGTCTGTGGGCAACTAGTAACAGAATCGGTTGTGCCATTAAATTTGTGTCAATAACATGAACATCTGGGGGCG
 ATATGGCCCAAAGCTGTCTACCTGGTGTGCAATTAATCCCCAAAGGGAACTGGTGGGGCCATGCCCTTACAAA
 CATGGGCGGCCCTGTCTGCTTGGCCACCTAGTTTGGAGGGGGCTGTAGAGAAAACTGTGCTACAAAGAAGGG
 TCAGACAGGTATTATCCCCCTCGAGAAGAGGAAAAAATAAGAAATAGAAGCAGCAGTCAAGTCCATGACACC
 CATGTCCGGACAAGATCAGATGATAGTAGCAGAAAAATGAAGTCATAAGCGCACAGCAATGTCCCAATTTGTTCT
 TGTGAAGTAAGATTAAAGAGATCAGTGCAAAAGGAAACAACTGCAATAGGTACGAATGTCTTGTGCTGGTGGT
 AGTAAGCTAAAGTTATTTGGCAGTGTACATATATGAATGCAATCCAGCATCTGTAGAGCTGCAATTCATTATGGT
 ATAATAGACAATGATGGTGGCTGGGTAGATATCACTAGACAAGGAAGAAAGCATTATTTTATCAAGTCCAAATAGA
 AATGGTATTCAAACAATTTGGCAAAATATCAGTCTGTCAATTTCTTCAAGTCTCTAAAGTAACAGTTTCAAGCTGTG
 ACTTGTGAACAACCTGTGGAAACAGCTCTGTCCATTTTCAATAAGCTGTCTTCAATTTGCCCAAGAGTATACGTCTCT
 CGTAACTGTATGCAAGCAAACTCCACATTATGCTCGTGAATTTGGAAGTCTGAGTTTATCTGATCTGTCCAGTATC
 TGCAGAGCAGCAGTACATGCTGGAGTGGTTCGAATACAGGTGGTGTATGTGTAGTGAATGCTGTGGACAAAAGA
 AAGACCTACATTGCTCTTTTCAAGATGGAATCTTCTCAGAAAGTTTACAGAACTCCTCAGGAGGAAAGGCAATC
 AGAGTGTGTGCTGTGTGTGGAACCTGAATACTTGAAGAGGACCATAAGACTATTCCAAATGCAATATTTCTGA
 ATTTTGTATAAACTGTAACATTACTGTACAGAGTACATCAACTATTTTTCAGCCCAAAAGGTGCCAAATGCATA
 TAAATCTTGATAAAACAAAGTCTATAAAATAAAACATGGGACATTAGCTTTGGGAAAAGTAATGAAAATATAATGG
 TTTTGAATAATCTGTGTTTAAATATTGCTATATTTTCTAGCAGTTATTTCTACAGTTAATTACATAGTCATGATT
 GTTCTACGTTTCATATATATATGGTGTCTTGTATATGCCACTAATAAATGAATCTAAACATTGAATGTGAATG
 GCCTCAGAAAAATCATCTAGTGCAATTTAAATAATATCGACTCTAAACTGAAGAAACCTTATCACTTTTCCCC
 AGTTCAATGCTATGCCATTACCAACTCCAAATAATCTCAATAATTTTCCACTTAATAACTGTAAAGTTTTTTC
 TGTAAATTTAGGCATATAGAATATTAATCTGATATTGCACTTCTTATTTTATATAAAATAATCCCTTTAATATC
 CAATGAATCTGTTAAAGTGTGTTGATTCCTTGGGAATGGCCTTAAAAATAAATGTAATAAAGTCAGAGTGGTGGT
 ATGAAACATTCCTAGTGATCATGTAGTAAATGTAGGGTTAAGCATGGACGCCAGAGCTTCTATGTACTGTATA
 AAATTGAGGTACATATTTCTTTTGTATCCTGGCAAAATCTCCTGCAGGCCAGGAAGTATAATAGCAAAAAGTT
 GAACAAAGATGAACATAATGTATTACATTACCATGCCACTGATTTTTTTTAAATGGTAAATGACCTTGTATATAA
 ATATTGCCATATCATGGTACCTATAATGGTGATATATTTGTTCTATGAAAAATGATTTGTGCTTGTATCTAAA
 AATCTGTAATGTTAGTTTGGTAAATTTTTTTCTGCTGGTGGATTACATATTAATTTTTTCTGCTGGTGGGA
 TAAACATTAAATTAATCATGTTTCAAAAAA

09570437.1.14.601

FIGURE 147

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45417

<subunit 1 of 1, 500 aa, 1 stop

<MW: 56888, pI: 8.53, NX(S/T): 2

MKCTAREWLRVTTVLFMARAIPAMVVPNATLLEKLLLEKYMDEDGEWWIAKQRGKRAITDNDM
QSILDLHNKLR SQVYPTASNMEYMTWDVELERSAESWAESCLWEHGPASLLPSIGQNLGAHW
GRYRPPTFHVQSWYDEVKDFSYPYEHECNPYCPFRCSGPVCTHYTQVVWATSNRIGCAINLC
HNMNIWGQIWPKAVYLV CNYSPKGNWWGHAPYKHGRPCSACPPSFGGGCRENL CYKEGSDRY
YPPREETNEIERQQSQVHDTHVTRSDSSRNEVIS AQQMSQIVSCEVRLRDQCKGTTTCNR
YECFAGCLDSKAKVIGSVHYEMQSSICRAAIHYGIIDNDGGWVDITRQGRKHYFIKSNRNGI
QTIGKYQSANSFTVSKVTVQAVTCETTVEQLCPFHKPASHCPRVYCPRNCMQANPHYARVIG
TRVYSDLSSICRAAVHAGVVRNHGGYVDVMPVDKRKTYIASFQNGIFSESLQNPPGGKA FRV
FAVV

Important features:

Signal peptide:

amino acids 1-20

Extracellular proteins SCP/Tpx-1/Ag5/PR-1/Sc7 protein

amino acids 165-186, 196-218, 134-146, 96-108 and 58-77

N-glycosylation site

amino acids 28-31

FIGURE 148

GCGGAGACAAGCGCAGAGCGCAGCGCACGGCCACAGACAGCCCTGGGCATCCACCGACGGCG
CAGCCGGAGCCAGCAGAGCCGGAAGGCGCGCCCGGGCAGAGAAAGCCGAGCAGAGCTGGGT
GGCGTCTCCGGGCGCCGCTCCGACGGGCAGCGCCCTCCCCATGTCCCTGCTCCCACGCCG
CGCCCTCCGGTCAGCATGAGGCTCCTGGCGGCCGCGCTGCTCCTGCTGCTGCTGGCGCTGT
ACACCGCGCGTGTGGACGGGTCAAATGCAAGTGCTCCCGGAAGGGACCCAAGATCCGCTAC
AGCGACGTGAAGAAGCTGGAATGAAGCCAAAGTACCCGCACTGCGAGGAGAAGATGGTTAT
CATCACCAACAAGAGCGTGTCCAGGTACCGAGGTGAGGAGCACTGCCTGCACCCCAAGCTGC
AGAGCACCAAGCGCTTCATCAAGTGGTACAACGCCTGGAACGAGAAGCGCAGGGTCTACGAA
GAATAGGGTGAAAAACCTCAGAAGGAAAACTCCAAACAGTTGGGAGACTTGTGCAAAGGA
CTTTGCAGATTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGCCTTTT
TTTCTCACAGGCATAAGACACAAATTATATATTGTTATGAAGCACTTTTTACCAACGGTCAG
TTTTTACATTTTATAGCTGCGTGCGAAAGGCTTCAGATGGGAGACCCATCTCTCTTGTGCT
CCAGACTTCATCACAGGCTGCTTTTTTATCAAAAAGGGGAAAACTCATGCCTTTCTTTTAA
AAAATGCTTTTTTGTATTTGTCCATACGTCACTATACATCTGAGCTTTATAAGCGCCCGGA
GGAACAATGAGCTTGGTGGACACATTTCAATGCAGTGTTGCTCCATTCTAGCTTGGGAAGC
TTCCGCTTAGAGGTCTTGCGCCTCGGCACAGCTGCCACGGGCTCTCTGGGCTTATGGCCG
GTCACAGCCTCAGTGTGACTCCACAGTGGCCCTGTAGCCGGGCAAGCAGGAGCAGGTCTCT
CTGCATCTGTTCTCTGAGGAACTCAAGTTTGGTTGCCAGAAAAATGTGCTTCATTCCTCCCT
GGTTAATTTTTACACACCCTAGGAAACATTTCCAAGATCCTGTGATGGCGAGACAAATGATC
CTTAAAGAAGGTGTGGGGTCTTTCCCAACCTGAGGATTTCTGAAAGGTTACAGGTTCAATA
TTAATGTCTCAGAAGCATGTGAGGTTCCCAACACTGTGACGAAAAACCTTAGGAGAAAACT
TAAAAATATATGAATACATGCGCAATACACAGCTACAGACACACATTCTGTTGACAAGGGAA
AACCTTCAAAGCATGTTTCTTTCCCTCACCACAACAGAACATGCAGTACTAAAGCAATATAT
TTGTGATTCCCCATGTAATCTTCAATGTTAAACAGTGCAGTCCTCTTTCGAAAGCTAAGAT
GACCATGCGCCCTTTCCTCTGTACATATACCTTAAAGAACGCCCCCTCCACACACTGCCCC
CAGTATATGCCGATTGTACTGCTGTGTTATATGCTATGTACATGTCAGAAACCATTAGCAT
TGCATGCAGGTTTCATATTTCTTCTAAGATGGAAAGTAATAAAATATATTTGAAATGTAAAA
AAAAAAAAAA

FIGURE 149

MSLLPRAPPVSMRLAAALLLLLALYTARVDGSKCKCSRKGPKIRYSDVKLEMKPKYPH
CEEKMWIITTKSVSRYRGQEHCLHPKLQSTKRFIKWYNWNEKRRVYEE

Signal sequence:

amino acids 1-34

[illegible]

GCCCAGGGAGCTGCTATGGCTTCCTTTGTGTGTACCCCGGCTCTGCGCTCATGTTAAACTCCAATGTCTCTCTGTG
 TAAATCTGCTCTTGCCATCAAGTTTACCCCTCATTGACAGCCAAGCAGAGTATCCAGTTGTCAACACAATTAATGG
 CAAATCCGGGGCCCTAAGAAACCGCTTACCCAAAGAGATCTTGGGTCCAGTGGCAGAGCTATAGGGGTCCCCA
 TGCTCCACCCCCCATCTGGAGAGGGCGGTTTCAGCCGCCAAGCCCGCTCTCTCGATAGTGGCATCCGAATATAC
 TACTCAGTTTGTCTGTGTGTCGCCCAGCAGTGGATGAGAGATCCTTACTGATGACATGCTGCCCATCTGTTT
 TGCCGCCAATTGGATATCTTTGATGACCTATGTTCAAGATCAAAATGAAGACTGCTTTTCTTAAACATCTAGT
 TCCCGCGGAAGATGGAGCCAAACAAGAAACACGAGCATGATATAACGAGTAATGACCGTGGTGGAGACGAAGA
 TATTCATGATCAGAACGTAAGAAGCCCGTCATGGTCTATATCCATGGGGATCTTACATGGAGGSCACCGSCA
 CATGATTGACCGGCAGCATTTTGGCAGGCTACGGAAAGCATGTCGTATACCAATTAACCTCCGTCTGGGAATCT
 AGGGTTTAAAGTACCGGTGACAGGACAAAGGCAATCTAGGCTCTCGTATGAGTCAAGCATCAGCATCGCGTG
 GATTGAGGAGAAATGTGGGAGCGCTTTGGCGGGACCCCAAGAGATGACCATCTTTGGCTCGGGGCTCGGGGCTC
 CTGTGTGACCGCTGTTTGACCTCTGTTCAGAAAGTCTCTTCCAGAGGCCATTAITCAGAGCGGCACCGC
 CTTGTCCAGCTGGGCAGTGAATACACGACCGCAAGTACATCGGATATGGCAGCAAGGTGGCTGGCAACAT
 CTGGACACCCAGCGCATGTAGATGCTTCGGCAACAGAACTCAAGAGCATCATCCAGCAGCATCACTCCCG
 GGCACCTTACCATAGCCTTCGGGCGCGTGATCGACGGCGACGTATCCAGACGACCCCGAGATCCTGATGGA
 CGTGGCGAGTGTCTTCAACTACGACATCATGCTGGGCTCAACCAAGGGGAAGGCGCTGAAGTTCTGTGGACGGCAT
 CTTGGATATCAGAGGACGGGTGTGACGCCAACGACATTTGACTTCTCGTGTCCAATCTGTGGACAACCTTTAGCG
 CTACCTCTGAAGGGAAAGACATCTTGGCGGAGATCATCAAGTTCTATGTACACAGATGGGCGGATGAAGAAAACCC
 GGAGCGCGCGGAAACACCTGTGTGGCTCTCTTTACGACACCAAGTGGGTGGCCCGCGCTGGCCGCGCACTCT
 GACGCGCAGTACCGCTCCCCACCTACTTCTATGCTCTTATCATCATCTGCCAAAGCGAATAGGACCGCACTG
 GCGAGATTGGTCCCATCTGTGTAGGTGGCCCTATGTTCTTCGCGATCCCCATCTCGTCCGACCGAGCTCTTAC
 TTGTAACCTTTTCCAGAAGCAGCATCTAGCTCAGCGCCGTGGTCTATGACCTACTGGACGAATCTCGCCAAATGCTG
 TGATCCAATTAACCAAGTTCTCAGGATACCAAGTTTATTCAACAACCAACCGCTTTGAAGAAGTGGCTGG
 CTCGAAGATAATCCCAAGACAGCTCTATCTGTCATATGGCTTGAACCAGAGTGAGAGATCACTACCGGGC
 AAGCAAGGTGGCTTTCTGGTTGGAACCTCGTTCTCATTTTGACAACTTGAACGGGATATCCAGTATGTTTCAAC
 AACCACAAGGTTTCTCCACGACATGACATCATTTCCCTATGGCACCGGAGATCTCCGCGCAAGATATGGCC
 AACCAACAAACCGCAGCAATCATCTCGCCAACTACCCAAACATCTATGAAGACCTCAACAACAGGCGCTG
 GGACACACATCTGCTCTATTGAACCAACAGGATATTCCACCGAATTAAAGTCTCAGCTTCGCGTGGGCGCTG
 GCTCCTCTTCCCAACATCTTAGCTTTTGGCGCGCTGTACTACAAAGAGCAAGAGGCGCATGAGATCTCAG
 GCGCCGATCTCCCGAGAAACACCAAAATGATCTCGTCACATCCAGAACGAAGAGATCTGCTCTGCAAT
 GAGACGCTGGAAACAGATCAGAGTGTGAGTCTGTGAGGCAAGACATCACTAGGCTCACTGCGCGCAGA
 CTACACCTCTCAGCTGCGCGCGTGCAGAGATGACATCCCATTTAGCAGCCAAACACCATCACTGATTTCCAA
 CACATGACGGGGATGACGCTTTGCCACATCTTTAACTCTTCTAGTGAGGAGACAAACAGTACAATTTACCCA
 CGGACATTTCCACCAATAGAGTATAGCTTTAGCGCTATTTACGCTCTATTCCTCTGCCCCTCCGCTCCGCTCAGCAAC
 AGAAGAGGAGGAGAAAGAGAGAAAGGAGAGAGAGAGAAAGAGTCTCCAGACAGGAATGTTTTGTGCCATCT
 GACTTAAAGACAAAAATGCAAAAAGGCAGTCATCCCATCCCGCGAGACCTTATCTGTTGTGTTTTCAGATTTAC
 AAGATCAACTCTTGACCTCTGTAATGTGAGAAGTACACATTTCTGTAAAAATCACTGTTTAAAGATCTTACCA
 CTCATTAATGTTTTAGTGTGATAGGACATCACATTTCAAGCCCGCGGTGTTTCCACGCTCATGGAGACGCT
 GACACTCTGAAATCTCCCAAGACATCTGATATTTTAAATTCATATGGAAGTTTAAACATTTCTTCTGTGTC
 CACAATGAGTGGCTCTGCTTAAGTGAAGAAAGAGTCAATGAGATTCTCCAGACATGAGACTGTAATCTCAG
 AGAGAGGAAACCTAGAATTTTATTTTAAAGAAATGGACTGTGCAGCGAAATCTGTACGGTTCTGTGCAAGAAG
 GTGTTTGTTCAGCTCGAACTATTTTAAAGAGACTTTGT

FIGURE 151

MLNSNVLLWLTALAIKFTLIDSQAQYPVVNTNYGKIRGLRTPLPNEILGPVEQYLGVPYASP
PTGERRFQPPPEPSSWTGIRNTTQFAAVCPQHLDERSLLHDMPLIWFTANLDTLMTYVQDQN
EDCLYLNIIYVPTEDGANTKKNADDITSNDRGEDEDIHDQNSKKPVMVYIHGGSYMEGTGNMI
DGSILASYGNVIVITINYRLGILGFLSTGDQAAKGNYGLLDQIQALRWIEENVGAFGGDPKR
VTIFGSGAGASCVSLLTLTSHYSEGLFQKAIQSGTALSSWAVNYQPAKYTRILADKVGCNML
DTTDMVECLRNKNYKELIQQTITPATYHIAFGPVIDGDVIPDDPQILMEQGEFLNYDIMLGV
NQGEGLKFVDGIVDNEDGVTPNDFDFSVSNFVDNLYGYPEGKDTLRETIKFMYTDWADKENP
ETRRKTLVALFTDHQWVAPAVAADLHAQYGSPTYFYAFYHHCQSEMKPWSADSAHGDEVFVY
FGIPMIGPTLFCSCNFSKNDVMLS AVVMTYWTNFARTGDPNQPVQDTKFIHTKPNRFEEVA
WSKYNPKDQLYLHIGLKPRVRDHYRATKVAFWLELVPHLHNLNEIFQYVSTTTKVPPDMS
FPYGTRRSPAKIWPTTKRPAITPANNPKHSKDPHKTGPEDTTVLIETKRDYSTELSVTI AVG
ASLLFLNILAFAALYYKKDKRRHETHRRPSPQRNTTNDIAHIQNEEIMSLQMKQLEHDECE
SLQAHDTLRLTCCPDYTLTLRRSPDDIPLMTPNTITMIPNTLTGMQLHTFNTFSGGQNSTN
LPHGHSTTRV

Signal sequence:

amino acids 1-24

Transmembrane domains:

amino acids 189-204, 675-692

FIGURE 153

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50911

><subunit 1 of 1, 348 aa, 1 stop

><MW: 39711, pI: 8.70, NX(S/T): 1

MAATLGPLGSWQQWRRCLSA RDGSRMLLLLLLLGSGQGPQVGAGQTFEY LKREHSLSKPYQ
GVGTGSSSLWNL MGNAMVMTQYIRLTPDMQSKQGALWNRVPCFLRDWELQVHF KIHGQGKKN
LHGDGLAIWYTKDRMQPGPVFGNMDKFVGLGVFVD TYPNEEKQQERVFPYISAMVNNGSLSY
DHERDGRPTELGGCTAIVRN LH YDTFLVIRYVKRHLTIMMDIDGKHEWRDCIEVPGVRLPRG
YYFGTSSITGDLSDNHDVISLKL FELTVERTPEEEKLHRDVFLPSVDNMKLPENTAPLPPLS
GLALFLIVFFSLVFSVFAIVIGIILYNKWQEQSRKRFY

Signal sequence:

amino acids 1-38

Transmembrane domain:

amino acids 310-329

00070400-101000

FIGURE 155

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48329

><subunit 1 of 1, 480 aa, 1 stop

><MW: 55240, pI: 9.30, NX(S/T): 2

MLFRNRFLLLLALALLAFVSLSLQFFHLIPVSTPKNGMSSKSRKRIMPDPVTEPPVTDPVY
EALLYCNIPSVASERSMEGHAPHHFKLVSVHVFIRHGDRYPLYVIPKTKRPEIDCTLVANRKP
YHPKLEAFISHMSKSGSGASFESPLNSLPLYPNHPLCEMGELTQTGTGVVQHLQNGQLLRDIYLK
KHKLLPNDWSADQLYLETTGKSRTLQSGLALLYGFLPDFDWKKIYFRHQPSALFCSGSCYCP
VRNQYLEKEQRRQYLLRLKNSQLEKTYGEMAKIVDVPTKQLRAANPIDSMMLCHFCHNVSFPC
TRNGCVDMEHFKVIKTHQIEDERERREKKLYFGYSLGHPILNQITGRMQRATEGRKEELF
ALYSAHDVTLSPVLSALGLSEARFPRFAARLIFELWQDREKPSHSVRILYNGVDVTFHTSF
CQDHHKRSPKPMCPLENLVRVFKRDMFVALGGSGTNYDACHREGF

Signal sequence:

amino acids 1-18

FIGURE 157

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48306

><subunit 1 of 1, 916 aa, 1 stop

><MW: 100204, pI: 4.92, NX(S/T): 4

MIPARLHRDYKGLVLLGILLGTLWETGCTQIRYSVP EELEKGSRVGDISRDLGLEPRELAER
GVRIIPRGRTQLFALNPRSGSLVTAGRIDREELCMGAIKCQLNLDILMEDKVKIYGVEVEVR
DINDNAPYFRESELEIKISENAATEMRFP LPHAWDPDIGKNSLQSYELSPNTHFSLIVQNGA
DGSKYPELV LKRALDREEKAAHHLVLTASDGGDPVRTGTARIRVMVLDANDNAPAFAPQPEYR
ASVPENLALGTQLLVNATDPDEGVNAEVRYSF RYVDDKAAQVFKLDCNSGTISTIGELDHE
ESGFYQMEVQAMDNAGYSARAKVLITVLDVNDNAPEVVLTSLASSVPENSPRGTLIALLNVN
DQDSEENGQVICFIQGNLPFKLEKSYGNYYSLVTDIVLDREQVPSYNITVTATDRGTPPLST
ETHISLNVADTNDNPPVFPQASYSAYIPENNPRGVSLVSVTAHDPDCEENAQITYSLAENTI
QGASLSSYVSINSDTGVLYALSSFDYEQFRDLQVKVMARDNGHPPLSSNVLSLFLVLDQNDN
APEILYPALPTD GSTGVELAPRSAEPGYLVTKVVAVDRDSGQNAWLSYRLLKASEPGLFSVG
LHTGEVRTARALLDRDALKQSLVVAVQDHGQPPLSATVTTLTVAVADSIPQVLADLGSLES PA
NSETSDLTLYLVVAVAAVSCVFLAFVILLALRLRRWHKSRLLQASGGGLTGAPASHFVGVD
GVQAFLQTY SHEVSLTTDSRKSHLIFPQPNYADMLVSQESFEKSEPLLLSGDSVFSKDSHGL
IEVSLYQIFFLFFFNCSVSQAGVQRYDHSSLRPQTPRLKQLSHLCLRCNRD YRCKPPTVCLS
IYLSIYLSIYLSIYLLLSCTDGS LTPVIPVLWEAEAGGSPEVGS LRPA

Signal sequence:

amino acids 1-30

Transmembrane domains:

amino acids 693-711, 809-823, 869-888

CCGAGGCTCTAGTGCAGGAGGAGAAGGAGGAGGAGCAGGAGGTGGAGATTCCAGTTAAAAAG
GCTCCAGAATCGTGTACACAGGCAGAGAACTGAAGTACTGGGGCTCTCCACTGGGTCGAA
TCAGTAGGTGACCCCGCCCTGGATTCTGGAAGACCTCACCATGGGACGCCCCGACCTCGT
GCGGCCAAGACGTGGATGTTCTCTGCTCTTGCTGGGGGGAGCCTGGGCAGGACACTCCAGGGC
ACAGGAGGACAAGGTGCTGGGGGGTCATGAGTGCCAACCCCATTCGCAGCCTTGGCAGGCGG
CCTTGTTCCAGGGCCAGCAACTACTCTGTGGCGGTGTCCTTGTAGGTGGCAACTGGGTCCTT
ACAGCTGCCCACTGTAAAAAACGAAATACACAGTACGCCTGGGAGACCACAGCCTACAGAA
TAAAGATGGCCCAGAGCAAGAAATACCTGTGGTTTCAGTCCATCCCACACCCCTGCTACAACA
GCAGCGATGTGGAGGACCAACCATGATCTGATGCTTCTTCAACTGCGTGACCAGGCATCC
CTGGGGTCCAAGTGAAGCCCATCAGCCTGGCAGATCATTGACCCAGCCTGGCCAGAAGTG
CACCGTCTCAGGCTGGGGCACTGTCACCAAGTCCCCGAGAGAAATTTCTGACACTCTCAACT
GTGCAGAAGTAAAAATCTTTCCCCAGAAGAAGTGTGAGGATGCTTACCCGGGGCGAGATCACA
GATGGCATGGTCTGTGCAGGCAGCAGCAAGGGGCTGACACGTGCCAGGGCGATTCTGGAGG
CCCCCTGGTGTGTGATGGTGCACTCCAGGGCATCACATCTGGGGCTCAGACCCCTGTGGGA
GGTCCGACAAACCTGGCGTCTATACCAACATCTGCCGCTACCTGGACTGGATCAAGAAGATC
ATAGGCAGCAAGGGCTGAATTCTAGGATAAGCACTAGATCTCCCTTAATAAACTCACAACTCT
CTGGTTC

FIGURE 159

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48336

<subunit 1 of 1, 260 aa, 1 stop

<MW: 28048, pI: 7.87, NX(S/T): 1

MGRPRPRAAKTWMFLLLLGGAWAGHSRAQEDKVLGGHECQPHSQPWQAALFQGQQLLCGGVL
VGGNWLTAAHCKKPKYTVRLGDHSLQNKDGPEQEIPVVQSIPHPCYNSSDVEDHNHDLMLL
QLRDQASLGSKVKPISLADHCTQPGQKCTVSGWGTVTSPRENFPDTLNCAEVKIFPQKKCED
AYPGQITDGMVCAGSSKGADTCQGDSSGGLVCDGALQGITSWGSDFCGRSDKPGVYTNICRY
LDWIKKIIGSKG

Important Features:

Signal peptide:

amino acids 1-23

Transmembrane domain:

amino acids 51-71

N-glycosylation site.

amino acids 110-113

Serine proteases, trypsin family, histidine active site.

amino acids 69-74 and 207-217

Tyrosine kinase phosphorylation site.

amino acids 182-188

Kringle domain proteins motif

amino acids 205-217

[illegible]

FIGURE 161

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA44184

><subunit 1 of 1, 473 aa, 1 stop

><MW: 50708, pI: 9.28, NX(S/T): 6

MKRASAGGSRLLAWLWLQAWQVAAPCPGACVCYNEPKVTTSCPQQGLQAVPVGIPAASQRI
FLHGNRISHVPAASFRACRNLTILWLHSNVLARIDAAFTGLALLEQLDLSDNAQLRSVDPA
TFHGLGRLHTLHLDRCLQELGPGGLFRGLAALQYLYLQDNALQALEPDDTFRDLGNLTHLFLH
GNRISSVPERAFRGLHSLDRLLHLHQNVRVAHVHPHAFRDLGRLMTLYLFANNLSALPTEALAP
LRALQYLRLNDNPWWCDRCRAPLWAWLQKFRGSSSEVP CSLPQRLAGRDLKRLAANDLQGCA
VATGPYHPIWTGRATDEEPLGLPKCCQPDAAKASVLEPGRPASAGNALKGRVPPGDSPPGN
GSGPRHINDSPFGTLPGSAEPPLTAVRPEGSEPPGFPTSGPRRRPGCSRKNRTRSHCRLGQA
GSGGGGTGDSESGALPSLTCSLTPLGLALVLWTVLGPC

Important features:

Signal peptide:

amino acids 1-26

Leucine zipper pattern.

amino acids 135-156

Glycosaminoglycan attachment site.

amino acids 436-439

N-glycosylation site.

amino acids 82-85, 179-183, 237-240, 372-375 and 423-426

VWFC domain

amino acids 411-425

FIGURE 162

GGAAGTCCACGGGGAGCTTGGATGCCAAAGGGAGGACGGCTGGGTCTCTGGAGAGGACTAC
 TCACTGGCATATTTCTGAGGTATCTGTAGAATAACCACAGCCTCAGATACCTGGGGACTTTAC
 AGTCCACAGAACCGTCTCTCCAGGAAGCTGAATCCAGCAAGAACAATGGAGGCCAGCGGGA
 AGCTCATTTCGAGACAAAGGCAAGTCCTTTTTCTTTCTCTTTTGGGGCTTATCTCTGGCG
 GCGCGCGCGGAACCTAGAAGCTATTCTGTGTGGTGGAGGAAATCTCCAGGCGGGGGTTAGGGTTG
 CAATTTAGCAAAAGGACTTGGGTCTGGAGCAGAGGGAAATCTCCAGGCGGGGGTTAGGGTTG
 TTTCCAGAGGGAACAACTACATTTCGAGCTCAATCAGGAGACCGCGGATTGTGTGCTAAAT
 GAGAAATTGGACCGTGAGGATCTGTGCGGTCAACAGAGCCCTGTGTGCTACGTTTCCAAGT
 GTTGCTAGAGAGTCCCTTCGAGTTTTTTCAAGCTGAGCTGCAAGTAATAGACATAAACGCAC
 ACTCTCCAGTATTTCTGGACAAACAAATGTTGGTGAAAGTATCAGAGAGCAGTCCTCTGGG
 ACTACGTTTTCTCTGAAGAATGCCGAAGACTTAGATGTAGGCCAAAACAATATTGAGAACTA
 TATAATCAGCCCCAACTCCTATTTTCGGGTCTCACCCGCAACCGCAGTGATGGCAGGAAAT
 ACCCAGAGCTGGTCTGGACAAAGCGCTGGACCGAGAGGAAGAAGCTGAGCTCAGGTTAACA
 CTCACAGCAGCTGGATGGTGGCTCTCCGCCCAGATCTGGCAGCTGCTCAGGTCTACATCGAAGT
 CCTGGATGTCAACGATAATGCCCTGAAATTTGAGCAGCCTTTCTATAGAGTCAGATCTCTG
 AGGACAGTCGCGTAGGCTTCTGTGTTGTGAAGGTCTCTGCCACGGATGTAGACACAGGAGTC
 AACGGAGAGATTTCTTATTCACTTTTCCAAGCTTCAGAAGAGATTGGCAAAACCTTTAAGAT
 CAATCCCTTGACAGGAGAAATTGAACTAAAAACAACCTCGATTTGAAAACTTCAGTCCT
 ATGAAGTCAATATTGAGGCAAGAGATGCTGGAACTTTTCGGAAAATGCACCGTTCTGATT
 CAAGTGATAGATGTGAACGACCATGCCCCAGAAGTTACCATGCTGCTGCTTACCAGCCCAAT
 ACCTGAGAACGCGCTGAACTGTGGTTGCACTTTTCAGTGTTCAGATCTTGATTGATTGAGGAG
 AAAATGGGAAAATTAGTTGCTCCATTGAGGAGGATCTACCCCTCTCTGAAATCCGCGGAA
 AACTTTTACACCTACTAACGGAGAGACCCTAGACAGAGAAAGCAGAGCGGAATACAACT
 CACTATCACTGTCACTGACTTGGGGACCCCTATGCTGATAACACAGCTCAATATGACCGTGC
 TGATCGCCGATGTCAATGACAACGCTCCCGCTTCACCCAAACCTCTACACCTGTTCGTC
 CGCGAGAACAAACAGCCCCGCTTGCACATCCGAGCGTCAAGCGCTACAGACAGAGACTCAGG
 CACCAACGCCAGGTCACCTACTCGCTGCTGCCGCCCGAGGACCCGACCTGCCCTCACAT
 CCCTGGTCTCCATCAACGCGGACAACCGCCACCTGTTGCCCCCAGGTCTCTGGACTACGAG
 GCCCTGCAGGGGTTCCAGTTCCGCGTGGGCGCTTCAGACCAAGCTCCCCGCGCTGAGCAG
 CGAGGCGCTGGTGCCTGTGTGTGTGCTGACGCGCAACGACAACCTCGCCCTTCGTGCTGTACC
 CGCTGCAGAACGGCTCCGCGCCCTGCACCGAGCTGGTGCCCCGGGCGGCGAGCCGGGCTAC
 CTGTGATCAAGGTGTGTGGCGGTGGACGGCGACTCGGGCCAGAACGCTGGCTGCTGTACCA
 GCTGCTCAAGGCCACGGAGCTCGGTCTGTTTGGCGTGTGGGCGCACAATGGCGAGGTGCGCA
 CCGCCAGGCTGCTGAGCGAGCGCGCAGCGGCCAAGCACAGGCTGGTGGTGTGTTCAAGGAC
 AATGGCGAGCCTCCGCGCTCGGCCACCGCCACGCTGCACGCTCCTGGTGGACGGCTTCTC
 CCAGCCCTACCTGCCTCTCCCGGAGGCGGCCCGACCCAGGCCAGGCCGACTTGTCTACCG
 TCTACCTGTGTGGTGGCGTTGGCCTCGGTGTCTTCGCTCTTCTCTTTTCGGTGTCTCTGTTC
 GTGGCGGTGCGGCTGTGTAGGAGGAGCAGGGCGGCCCTCGGTGGTCTGCTGTGGTGGCGGA
 GGGCCCCCTTCCAGGCGATCTTGTGGACATGAGCGGCACAGGACCTTATCCAGAGCTACC
 AGTATGAGGTGTGTCTGGCAGGAGGCTCAGGGACCAATGAGTTCAAGTTCTGAAGCCGATT
 ATCCCCAATTCCTTCCCCAGTGCCCTGGGAAAGAAATACAAGGAAATTCTACCTTCCCCAA
 TAACCTTGGGTTCAATATTCACTGACCATAGTTGACTTTTACATTCCATAGGATTTTATT
 TGTGGCATTTCATGCCAATGTTTATTTCGCCCAATTTGTGTGTATGTAATATTGTACGGAT
 TTACTCTTGATTTTTCTCATGTTCTTCTCCCTTTGTTTAAAGTGAACATTACCTTTATT
 CCTGGTCTCT

FIGURE 163

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48314

<subunit 1 of 1, 798 aa, 1 stop

<MW: 87552, pI: 4.84, NX(S/T): 5

MEASGKLICRQRQVLFSFLLGLSLAGAAEPRSYSVVEETEGSSFVTNLAKDLGLEQREFSR
RGVRVVSARGNKLHLQLNQETADLLLNEKLDREDLCGHTEPCVLRFRQVLLESPPFEFFQAELOV
IDINDHSPVFLDKQMLVKVSESSPPGTTTFPLKNAEDLDVGQNNIENYIISPNSYFRVLTRKR
SDGRKYPELVLDKALDREEEAELRLTLTALDGGSPPRSGTAQVYIEVLVDVNDNAPEFEQPFY
RVQISEDSPVGFLVVKVSATDVTGVNGEISYSLFQASEEIGKTFKINPLTGEIELKKQLDF
EKLQSYEVNIEARDAGTFSGKCTVLIQVIDVNDHAPVETMSAFTSPIENAPETVVALFSVS
DLDSGENGKISCISIQEDLPFLKSAENFYTLTTERPLDRESRAEYNITITVTDLGTPLMITQ
LNMTVLIADVNDNAPAFQTSTYTLFVRENNSPALHIRSVSATDRDSGTNAQVTYSLLPPQDP
HLPLTSLVSINADNGHLFALRSLDYEALQGFQFRVGASDHGSPALSSEALVRVVLDANDNS
PFVLYPLQNGSAPCTELVPRAAEPGYLVTKVAVDGDGQNAWLSYQLLKATELGFLGVWAH
NGEVRTARLLSERDAAKHRLVVLVKDNGEPPRSATATLHVLLVDGFSQPYLPLPEAAPTQAO
ADLLTVYLVVALASVSSLFVSVLLFVAVRLCRRSRAASVGRCLVPEGPLPGHLVDMGSTRT
LSQSYQYEVCLAGGSGTNEFKFLKPIIPNFPPQCPGKEIQGNSTFPNNFGFNIQ

Important features:

Signal peptide:

amino acids 1-26

Transmembrane domain:

amino acids 685-712

Cadherins extracellular repeated domain signature.

amino acids 122-132, 231-241, 336-346, 439-449 and 549-559

ATP/GTP-binding site motif A (P-loop).

amino acids 285-292

N-glycosylation site.

amino acids 418-421, 436-439, 567-570 and 786-789

ACCACACGCGTCCGCCACACGCGTCCGCCACGCGTCCGCCACGCGTCCGCGCGTAGCCGTGC
 GCCGATTGCTCTCGGCCTGGGCAATGGTCCCGGCTGCCGGTCGACGACGCCCCCGCGTCAT
 CGGGCTCCTCGGCTGGTGGCAAGTATTGCTGTGGGTGCTGGGACTTCCCGTCCGCGGCGTGG
 AGGTGTCAGAGGAAAGTGGTCGCTTATGGTCAGAGGAGCAGCCTGCTCACCTCTCCAGGTG
 GGGGCTGTGTACCTGGGTGAGGAGGAGCTCCTGCATGACCCGATGGGCCAGGACAGGGCAGC
 AGAAGAGGCCAATGCGGTGCTGGGGCTGGACACCCAAGGCGATCACATGGTGATGCTGTCTG
 TGATTCTCTGGGGAAGCTGAGGACAAAGTGAGTTCAGAGCCTAGCGCGTCACTGTGGTGCT
 GGAGGAGCGGAGGACTCAAGGTGCAACGTCGAGAGAGCCTTTTCTCTCTGGATGGCGCTGG
 AGCACACTTCCCTGACAGAGAAGAGGAGTATTACACAGAGCCAGAAGTGGCGGAATCTGACG
 CAGCCCCGACAGAGGACTCCAATAACACTGAAAGTCTGAAATCCCCAAAGGTGAACCTGTGAG
 GAGAGAAACATTACAGGATTAGAAAATTTCACTCTGAAAATTTTAAATATGTACAGGACCT
 TATGGATTTTCTGAACCCAAACGGTAGTGA CTGTACTCTAGTCTGTTTACACCCCGTGGT
 GCCGCTTTTCTGCCAGTTTGGCCCTCACTTTAACTCTCTGCCCGGGCATTTCAGCTCTT
 CACTTTTGGCACTGGATGCATCTCAGCACAGCAGCCTTCTACCAGGTTTGGCACCGTAGC
 TGTTCTCTAATATTTTATTATTTCAAGGAGCTAAACCAATGGCCAGATTAAATCATACAGATC
 GAACACTGGAACACTGAAAATCTTCATTTTAAATCAGACAGGTATAGAAGCCAAGAAGAAT
 GTGGTGGTAACTCAAGCCGACCAAAATAGGCCCTCTCCAGCACTTTGATAAAAAGTGTGGA
 CTGGTGTCTGTATTTCTCTATCTTTTTAATTAGTTTTATTATGTATGCTACCATTGAA
 CTGAGAGTATTCGGTGGCTAATTCAGGACAAGAGCAGGAACATGTGGAGTAGTGATGGTCT
 GAAAGAAGTTGGAAAGAGGAACTTCAATCCTTCGTTTCAGAAATTAGTGCTACAGTTTCATA
 CATTTTCTCCAGTGACGTGTTGACTTGAAACTTCAGGCAGATTAAAAGAATCATTTGTTGAA
 CAACTGAATGTATAAAAAAATTATAAACTGGTGTTTTAACTAGTATTGCAATAAGCAAATGC
 AAAAATATTCAATAG

FIGURE 165

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48333

><subunit 1 of 1, 360 aa, 1 stop

><MW: 39885, pI: 4.79, NX(S/T): 7

MVPAAGRRPPRVMRLLGWQVLLWVLGLPVRGVEVAEESGRLWSEEQPAHPLQVGAVYLGEE
ELLHDPMGQDRAABEANAVALGLDTQGDHVMVLSVIPGEAEDKVSSEPSGVTCGAGGAEDSRC
NVRESLFLSLDGAGAHFPDREEEYYTEPEVAESDAAPTEDSNNTESLKSPKVNCEERNITGLE
NFTLKIILNMSQDLMDFLNPNNGSDCTLVLFYTPWCRFSASLAPHFNSLPRAFPAHFLALDAS
QHSSLSTRFGTVAVPNILLFQGAQPMARFNHTDRTLETLETKIFIFNQGTGIEAKKNVVVTQADQ
IGPLPSTLIKSVDWLLVFSLFFLISFIMYATIRTESIRWLIPGQEQEHVE

Important features:

Signal peptide:

amino acids 1-25

Transmembrane domain:

amino acids 321-340

Homologous region to dilsufide isomerase

amino acids 212-302

N-glycosylation site.

amino acids 165-168, 181-184, 187-190, 194-197, 206-209, 278-281
and 293-296

Thioredoxin domain

amino acids 211-227

FIGURE 166

CCCCGGCTCCGCTCCCTCTGCCCCCTCGGGGTCGCGCGCCACGATGCTGCAGGGCCCTGGCT
CGCTGCTGCTGCTCTTCTCTCGCCTCGCACTGCTGCCTGGGCTCGGCGCGCGGGCTCTTCCTC
TTTGGCCAGCCCCGACTTCTCCTACAAGCGCAGCAATTGCAAGCCCATCCCGGTCAACCTGCA
GCTGTGCCACGGCATCGAATACCAGAACATGCGGCTGCCCA^ΔCTGCTGGGCCACGAGACCA
TGAAGGAGGTGCTGGAGCAGGCCGCGCTTGGATCCCGCTGGTCATGAAGCAGTGCCACCCG
GACACCAAGAAGTTCTGTGCTCGCTCTTCGCCCCCGTCTGCCTCGATGACCTAGACGAGAC
CATCCAGCCATGCCACTCGCTCTGCGTGCAGGTGAAGGACCGCTGCGCCCCGGTCATGTCCG
CCTTCGGCTTCCCCTGGCCCGACATGCTTGAGTGCACCGTTTCCCCCAGGACAACGACCTT
TGCATCCCCCTCGCTAGCAGCGACCACCTCCTGCCAGCCACCGAGGAAGCTCCAAAGGTATG
TGAAGCCTGCAAAAATAAAAAATGATGATGACAACGACATAATGGAACGCTTTGTAAAAATG
ATTTTGCACTGAAAATAAAAGTGAAGGAGATAACCTACATCAACCGAGATACCAAAATCATC
CTGGAGACCAAGAGCAAGACCATTTACAAGCTGAACGGTGTGTCCGAAAGGGACCTGAAGAA
ATCGGTGCTGTGGCTCAAAGACAGCTTGCAGTGCACCTGTGAGGAGATGAACGACATCAACG
CGCCCTATCTGGT CATGGGACAGAAA CAGGTTGGGAGCTGGTGATCACCTCGGTGAAGCGG
TGGCAGAAGGGGCAGAGAGAGTTCAAGCGCATCTCCGCGAGCATCCGCAAGCTGCAGTGC^{TA}
^GTCCCCGGCATCCTGATGGCTCCGACAGGCCTGCTCCAGAGCACGGCTGACCATTCTGCTCC
GGGATCTCAGTCCCCGTTCCCCAAGCACTCCTAGCTGCTCCAGTCTCAGCCTGGGCAGCT
TCCCCCTGCCTTTTGCACGTTTGCAATCCCCAGCATTTCTGAGTTATAAGGCCACAGGAGTG
GATAGCTGTTTTACCTAAAGGAAAAGCCCCCGAATCTTGTAGAAATATTCAAAC^{TA}TA
AAATCATGAATATTTTAA

FIGURE 167

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50920

><subunit 1 of 1, 295 aa, 1 stop

><MW: 33518, pI: 7.74, NX(S/T): 0

MLQGPGSLLLLFLASHCCIGSARGLFLFGQPDFSYKRSNCKPIPVNLQLCHGIEYQNMRLPN
LLGHETMKEVLEQAGAWIPLVMKQCHPDTKKFLCSLFAPVCLDDLDETIQPCHSLCVQVKDR
CAPVMSAFGFPWPDMLECDRFPQDNDLCIPLASSDHLLPATEEAPKVCEACKNKNDNDNDIM
ETLCKNDFALKIKVKEITYINRDTKIILETKSKTIYKLVGVSEKDLKKSVLWLKDSLQCTCE
EMNDINAPYLVMGQKQGGLVITSVKRWQKGQREFKRISRIRKLQC

Important features:

Signal peptide:

amino acids 1-20

Cysteine rich domain, homologous to frizzled N terminus

amino acids 6-153

FIGURE 168

GTGGAGGCCGCCGACG**AT**GGCGGGGCCGACGGAGGCCGAGACGGGGTTGGCCGAGCCCCGGG
CCCTGTGCGCGCAGCGGGGCCACCGCACCTACGCGCGCCGCTGGGGTGTTCCTGCTCGCGATC
AGCCTGCTCAACTGCTCCAACGCCACGCTGTGGCTCAGCTTTGCACCTGTGGCTGACGTCAT
TGCTGAGGACTTGGTCTCTGTCCATGGAGCAGATCAACTGGCTGTCACTGGTCTACCTCGTGG
TATCCACCCCATTTGGCGTGGCGGCCATCTGGATCCTGGACTCCGTCGGGCTCCGTGCGGCG
ACCATCCTGGGTGCGTGGCTGAACCTTTGCCGGGAGTGTGCTACGCATGGTGCCCTGCATGGT
TGTGGGACCCAAAACCCATTTGCCTTCTCATGGGTGGCCAGACCTCTGTGCCCTTGCCC
AGAGCCTGGTCATCTTCTCTCCAGCCAAGCTGGCTGCCTTGTGGTTCCCAGAGCACCAGCGA
GCCACGGCCAACATGCTCGCCACCATGTGGAACCTCTGGGCGTCTTGTGGCCAATGTGCT
GTCCCTGTGCTGGTCAAGAAGGGTGAGGACATTCCGTTAATGCTCGGTGTCTATACCATCC
CTGCTGGCGTCGTCTGCCTGCTGTCCACCATCTGCCTGTGGGAGAGTGTGCCCCCCACCCCG
CCCTCTGCCGGGGCTGCCAGCTCCACCTCAGAGAAGTTCTGGATGGGCTCAAGCTGCAGCT
CATGTGGAACAAGGCCTATGTCATCTGGCTGTGTGCTTGGGGGGAATGATCGGGATCTCTG
CCAGCTTCTCAGCCCTCCTGGAGCAGATCCTCTGTGCAAGCGGCCACTCCAGTGGGTTTTCC
GGCCTCTGTGGCGCTCTCTTCATCACGTTTGGGATCCTGGGGGCACTGGCTCTCGGCCCTTA
TGTGGACCGACCAAGCACTTCACTGAGGCCACCAAGATTGGCCTGTGCCTGTTCTCTCTGG
CCTGCGTGCCCTTTGCCCTGGTGTCCCAGCTGCAGGGACAGACCTTGCCCTGGCTGCCACC
TGCTCGTGCTCGGGCTGTTTGGCTTCTCGGTGGGCCCCGTGGCCATGGAGTTGGCGGTGGA
GTGTTCTTCCCCTGGGGGAGGGGGCTGCCACAGGCATGATCTTTGTGCTGGGGCAGGCCG
AGGGAATACTCATATGCTGGCAATGACGGCACTGACTGTGCGACGCTCGGAGCCGTCCTTG
TCCACCTGCCAGCAGGGGGAGGATCCACTTGACTGGACAGTGTCTCTGCTGTGATGGCCGG
CCTGTGCACCTTCTTCAGCTGCATCCTGGCGGTCTTCTTCCACACCCCATACCGGCGCTGC
AGGCCGAGTCTGGGGAGCCCCCTCCACCCGTAACGCCGTGGGCGGCGCAGACTCAGGGCCG
GGTGTGGACCGAGGGGGAGCAGGAAGGGCTGGGGTCTGGGGCCAGCACGGCGACTCCGGA
GTGCACGGCGAGGGGGGCTCGCTAGAGGACCCAGAGGGCCCGGAGCCCCACCCAGCCT
GCCACCGAGCGACTCCCCGTGCGCAAGGCCAGCAGCCACCGACGCGCCTCCCGCCCCGGC
AGACTCGCAGGCAGGGTCCAAGCGTCCAGGTTTATTGACCCGGCTGGGTCTCACTCCTCCTT
CTCTCCCCGTGGGTGATCACG**TAG**CTGAGCGCTTGTAGTCCAGGTGCCCCCCACATCGA
TGGAGGCGAACTGGAAATCTGGTCCACCTGCGGGCGGGGGCGAAAGGGCTCCTTGCGGGCT
CCGGGAGCGAATTACAAGCGCGCACCTGAAAA

FIGURE 169

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50988

><subunit 1 of 1, 560 aa, 1 stop

><MW: 58427, pI: 6.86, NX(S/T): 2

MAGPTEAETGLAEPRALCAQRGHRTYARRWVFLLAISLLNCSNATLWLSFAPVADVIAEDLV
LSMEQINWLSLVYLVVSTPFGVAAIWILDSVGLRAATILGAWLNFAGSVLRMVPCMVVGTON
PFAFLMGQSLCALAQSLVIFSPAKLAALWFPFHQRATANMLATMSNPLGVLVANVLSPLV
KKGEDIPMLGVYTIPAGVVCLLSTICLWESVPPTPPSAGAASSTSEKFLDGLKLQLMWNKA
YVILAVCLGGMIGISASFALLEQILCASGHSSGFGSLCGALFITFGILGALALGPYVDRTK
HFTEATKIGLCFLSLACVPFALVSQLOGQTLALAATCSLLGLFGFSVGPVAMELAVECSFPV
GEAATGMIFVLGQAEIGILMLAMTALTVRREPSLSTCQQGEDPLDWTVSLLLMAGLCTFF
SCILAVFFHTPYRRLQAESGEPPSTRNAVGGADSGPGVDRGGAGRAGVLGPSTATPECTARG
ASLEDPRGPGSPHPACHRATPRAQGPAATDAPSRPCRLAGRVQASRFIDPAGSHSSFSSPWVIT

Important features:

Signal peptide:

amino acids 1-44

Transmembrane domains:

amino acids 61-79, 98-112, 126-146, 169-182, 201-215, 248-268,
280-300, 318-337, 341-357, 375-387, 420-441

N-glycosylation site.

amino acids 40-43 and 43-46

Glycosaminoglycan attachment site.

amino acids 468-471

FIGURE 170

GTCCACATCTCTGCTCAACTGGGTCAAGTCCCTCTTAGACAGAGCTCTTGTCATCATTTTGTCTGAAGTGGACCAAC
 TAGTTCCTCCAGTAGGGGGTCTCCCTTGGCAATCTTGATCGCGCTTTGGACATCTCAGATCGCTTCCAAATGAAGA
 TGCCCTTGCTCTGGGGCTCTGCTTGTTTTCATAATCATCTAACTATGGGACAAAGTTGTGCCCGCAGCTCTGGGGG
 AAGGAGCACGGGGCTGATCAAGCAATCAAGGAAACACTGGAGGACTTGCTCCAGCTTGCTGAAAGAACTCTAGTGGTT
 TCTGAATCTAGCCCACTTGCGGTGAAGCATGATGCAACTCTGCAACTCTGCTGGGGCTTTTGGGGCCAGGTGG
 CTACTTATTTCTTTTAGGGGATTCTCAGGAGGTGACCACTCTCAGCGTGAAATCAAGTGTACAGAGGAAGTGCC
 ATCTGGTACAGTGATCGGGAAGCTGTGCCAGGAACCTGGGCCGGGAGGAGAGCGGAGGCAAGCTGGGGCCGCTT
 CCAGCTGTGTGAGCTGCTCAGGCGCTCCCCATTGAGGTGACCTCTGAGGAAGGCTTGCTCAGCACAGGCGAGCG
 GCTGGATCAGAGAGCATGTGCGGACAGTGGGATCCTGCGCTGGTTTCTCTTGATGTGCTTGGCCAGGGGATTT
 GGCTCTGATCATGTGGAGATCCAAGTGTGGACATCAATGACCACAGCCACGCTTTCCTCAAGGCGAGGACGGA
 GCTGGAAATCTCTGAGAGCGCTCTCTGCAACCCGATCCCCCTGGACAGAGCTCTTGACCCAGACACAGGCCC
 TAACACCTTGACACCTTACACTCTGTCTCCAGTGAGCACTTGCTTGGATGTCTATGTGGCCCTGATGAGAC
 CAAACATGCAGAACTCATAGTGGTGAAGGAGCTGGACAGGAAATCCATTCTTTTGATCTGGTGTAACTGC
 CTATGACAATGGGAACCCCCCAAGTCAGGTACAGCTTGGTCAAGGTCAACGCTCTTGGACTCCAATGACAATAG
 CGCTGCGTTTGCTGAGAGTTTCACTGTCACCTGGAATCCAAAGAGATGCTGCACCTGGTACGCTTCTCATAAAAC
 GACCGCCACAGACCTGACCAAGGCCCAATGGGGAGGTGGAGTTCTTCTCAGTAAGCAATGCTCCAGAGGTT
 GCTGGACACCTTTCAGTATTGATGCGCAAGACAGGCCAGGTCTTCTGCTGCACTCTAGACTATGAGAAAACCC
 TGCCCTACGAGGTGGATGTTTACGCAAGGAGCTGGGTCCCAATCTATCCAGGCCATTTGCAAGTCTTCTCATCA
 GGTCTCGATGTCAATGACACATCCCAAGCATCCACGTACATGGGCTCCAGCCATCACTGGTGTGAGAAGC
 TCTTCCCAAGGACAGTTTATTTGCTCTTGTCTGTCAGATGACTTGGATTGAGGACAAATGGTTTGTGCTCACTG
 CTGCGTGAGCCAAAGAGCTGGGCCATTGAGGTGAAAGAACTAATGGCAACACATACATGTTGCTAAACCAATGC
 CACATCTGGACAGAGAGAGTGGGCCAAATATACCTTCACTCTGTTAGCCCAAGACCAAGGACTCCAGCCCTTATC
 AGCCAAAGAAACAGCTCAGCATTCAGATCAGTGACATCAACGACAAATGCACTCTGTTTGGAGAAAAGCAGGTATGA
 AGTCTCCACGCGGGAAAACAACTTACCCCTCTTCACTCTATACCATCAAGGCTCATGATGCAAGTCTGGGCAT
 TAATGAAAAGTCTCATACCCGATCCAGGACTCCCAAGTGGTCACTTATGATCTTGTACTCCACACAGGAGA
 GGTCACTGCTCAGAGGTCATGAACATGAGAGATGGCCGGCTTTGAGTTCCAGGTGATCGCAGAGGACAGCGG
 CCAGCCCATGCTTGATCCAGTGTCTCTGTGTGGGTGAGCTCTTGGATGCCAATGATAATGCCCCAGAGGTGGT
 CCAGCTGTGCTCAGCGATGGAAAAGCCAGGCCCTCTCGTGCTTGTGAATGCCCTCCAGGCCCACTTGCTGGTGGCC
 CATCAGAGCTCCCAATGGCTTGGGCCAGCGGGCACTGACACACCTCCACTGGCCCATCAGACTCCCGCCCAAT
 CTTTGTGACCAACTTGTGGCAAGATGCAGACTCGGGGCCAAATGGAGAGCCCTTCTCAGCATCCGCAATGG
 AAATGAAGCCCACTCTTCACTCTCAACCTCTATAGGGGCACTGTTGCTCAATGTGACCAATGGCAGCAGCGCT
 CATTTGGGAGTGAGTGGGAGCTGGAGATAGTAGTAGAGGAGCAGGGAAGCCCCCCCTTGACAGCCCGAGCCCTT
 GAGGCTCATGTTTCTCACCAGTGTGGACCACTTGAAGGACTCAGCCGCAAGCTTGGGCGCTTGAGCATGTGAT
 GCTGACGCTGATCTGCTGCTGTACTGTTGGGCATCTTCCGGTTGATCTTGGCTTTGTTTCACTGCTCATCTGCGG
 GACAGAAAAGAAAGGACAAAGGGCCCTACAATCTGTCGGGAGGCCAGTCCACTACCGCCAGCAGCCCAAGAGGCC
 CCAGAAACACATCTGACAGGCGAGACATCCACTCTGTCCTGTGCTCAGGGGTGAGCCAGGTGAGCCTTTGTGAAT
 CGCGGCTCCCTCAAAAGATGTGGACAAGGAGCGATGATGGAAGCAGGCTGGAGCCCTGCTCTGAGGCCCTCTT
 CCACCTTACCCCGACCTGTACAGGACGCTGCTAATCAAGCAACCAAGGAGCAACCGCGGAGAGCGGAGAGGT
 GCTGCAAGACAGGTCACCTCTCTTTCAACCATCCAGGACAGGAATGCTTCCCGGAGAACCTGAACCTTCC
 CGAGCCCAAGCTGCCACAGGCCAGCCAGTTCAGGCCCTCTGAAGTTTGCAGGAGCCGCCACAGGGAGGCTGGC
 TGGAGACAGGCGAGTGAGGAGGCCCAAGAGGCCACAGCCTCTCTGCAACCTTGAGACCGGACGCGACATCT
 CAATGGCAAGTGTCTCCCTGAGAAAAGATCAGGGCCCTGTCAGATCTTCCGAGAGCTTGTCCGCTGTCTGTGCT
 TGCTTCTGCGGAGCGGAACCCCTGGAGGAGCTCACTGTGATTTCTCTCTGTGTCAGCAAACTTCCAGCTGCT
 GTCTTGTGCTCAGGCGCAATCTGAGACCAATCAGATGGCCCAAGTGCAGGGCTGGAGGCCAGACAGACCCAGAACAGGAGGA
 CAGCAGGAGTGCAATCCGACAGGACAGGACGAGGCAAGCAGCAGCAGCAGCAGCTGCTGACACATCTCTTGGCCCTCAC
 AGGCGCTTGGATCTCTGAAGAGGACCTCTCTGTGAAGCAACTGTGAGAAAGAGCTGTCAAGTCTGTCTGGACCC
 CAGCAGAGTCTGGCCCTTGAAGCCGCTGAGCGCCCTTGAACCGGCTGATGGCGAGACTCTTTTGGCCCTCAC
 CACCAATACCTGACAAATGTGATCTCCCGGATGCTGACGCCAGCGAGGAGCCGAGGACCTTCCAGACGTTCCG
 CAAAGCAGAGGCACACAGCTAGGTGAGGCCCAAGCAGGACAGGCTTGGCCAGCAGCTTGTCTCGGAGATGAGCTCACT
 GCTGGAGATGCTGCTGGAACAGCGCTCCAGCATGCCCTGGAGGCCGCCCTCGAGGCGCTCGCGCGCTCTCTGCT
 CTGCGGGAGGACCTTCACTTTAGACTTGGCCACAGTGCAGCTCAGGCAATGAAAGTGCAGAGGGACCCAGGTGG
 AAGAGCGGGACTGAGGGCAAGGACGAGGCAAGCAGCAGCAGCAGCAGCTGCTGACACATCTCTCAGACGCTT
 CTGGATCCAGAAACAGGGGGCTGAGATCTGTGGAACAGAGCTGTTTCTAAATCTTGTAACTCTACTAGCTAG
 CGCGGCTCTGAGAACTTAGGTGACTTATGCTACCCCAAGAGGAGGAGGAGGAGCCCAAGCACTAGCAGCTGAC
 TGACCAAGCAGCCCTTGTGAAGCAGCTCTGAGTCTTTTGGAGGACAGGAGGCTTGTGCTGAGATAGTGT
 TCTTGGCAAAACATATGTGGAGCAAAAGGTGAGTCTCTTGGCAGAACAGTGCACCGAGATACAGGAGG
 AAGGTTGGCTCTTGTGGAGGAGGTGAGGGGCTGTACCTGGGGGTGCCAGGAAATGCTCTGACCTAT
 CAATAAAGGAAAGCAGTAAAAAATAAAAAAAAAA

09978403.101601

FIGURE 171

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48331

<subunit 1 of 1, 1184 aa, 1 stop

<MW: 129022, pI: 5.20, NX(S/T): 5

MMQLLQLLGLLGPGGYLFLGLGDCQEVTTLTVKYQVSEEVPSGTVIGKLSQELGREERRRQA
GAAFQVLQLPQALPIQVDSEEGLLSTGRRILDREQLCRQWDPCLVSFDVLATGDLLIHVEIQ
VLDINDHQPRFPKGEQELEISESASLRTRIPLDRALDPDTGPNLTHTYTLSPSEHFALDVIV
GPDETKHAELIVVKELDREIHSFFDLVLTAYDNGNPPKSGTSLVKVNVLDNSNDNSPAFAESS
LALAIQEDAAPGTLTIKLTATDPDQGPNGEVEFFLSKHPPEVLDTFSIDAKTGQVILRRPL
DYEKNPAYEVDVQARDLGPNPIPAHCKVLIKVLVDVNDNIPSIHVTWASQPSLVSEALPKDSF
IALVMADDLDLSDGHNLVHCWLSQELGHFRLKRTNGNTYMLLTNATLDREQWPKYTLTLTLLAQD
QGLQPLSAKKQLSIQISDINDNAPVFEKSRYEVSTRENNLPSLHLITIKAHADLGINGKVS
YRIQDSPVAHLVAIDSNTGEVTAQRSLNYEEMAGFEFQVIAEDSGQPMLASSVSVVWSLLDA
NDNAPEVVQPVLSDGKASLSVLVNASTGHLLVPIETPNGLGPAGTDTPLATHSSRPFLTT
IVARDADSGANGEPLYSIRNGNEAHLFILNPHTGQLFVNVNASSLIGSEWELEIIVVEDQGS
PPLQTRALLRVMFVTSVDHLRDSARKPGALSMMLTVICLAVLLGIFGLILALFMSICRTEK
KDNRAYNCREAESTYRQQPKRPQKHQKADIHLVPVLRGQAGEPCEVGQSHKDVDKEAMMEA
GWDPCQLQAPFHLTPTLYRTLNRNQGNGAPAESREVLQDTVNLLFNHPRQRNASRENLNLP
QPATGQPRSRPLKVAGSPTGRLAGDQGSEAPQRPASSATLRRQRHLNGKVSPEKESGPRQ
ILRSLVRLSVAFAERNPVBELTVDSPPVQQISQLLSLLHQQFQPKPNHRGNKYLAKPGGS
RSAIPDTDGP SARAGGQTDPEQEEGPLDPEEDLSVKQLLEBELSSLLDPSTGLALDRLSAPD
PAWMARLSLPLTTNYRDNVISPDAAATEEPRTFQTFGKAEAPELSPTGTRLASTFVSEMSSL
LEMLLEQRSSMPVEAASEALRRLSVCGRTLSDLATSAASGMKVQDPPGGKTGTGKSRGSS
SSSRCL

Important features:

Signal peptide:

amino acids 1-13

Transmembrane domain:

amino acids 719-739

N-glycosylation site.

amino acids 415-418, 582-585, 659-662, 662-665 and 857-860

Cadherins extracellular repeated domain signature.

amino acids 123-133, 232-242, 340-350, 448-458 and 553-563

FIGURE 172

CGGACGCGTGCGGACGCGTGGGGGAGAGCCGAGTCCCGGCTGCAGCACCTGGGAGAAGG
CAGACCGTGTGAGGGGGCCGTGGCCCCAGCGTGCTGTGGCTCTGGGGAGTGGGAAGTGGAG
GCAGGAGCCCTTCCTTACACTTCGCCATGAGTTTCTCATCGACTCCAGCATCATGATTACCT
CCCAGATACTATTTTTTGGATTGGGTGGCTTTTCTTCATGCGCAAATTGTTTAAAGACTAT
GAGATACGTCAGTATGTTGTACAGGTGATCTTCTCCGTGACGTTTGCATTTTCTTGACCAT
GTTTGAGCTCATCATCTTTGAAATCTTAGGAGTATTGAATAGCAGTCCCGTATTTTCACT
GGAAATGAACCTGTGTGTAATCTGCTGATCCTGGTTTTTCATGGTGCCTTTTTACATGGC
TATTTTATTGTGAGCAATATCCGACTACTGCATAAACAACGACTGCTTTTTTCTGTCTCTT
ATGGCTGACCTTTATGTATTTCTTCTGGAACTAGGAGATCCCTTTCCCATTTCTCAGCCAA
AACATGGGATCTTATCCATAGAACAGCTCATCAGCCGGGTGGTGTGATTGGAGTGACTCTC
ATGGCTCTTCTTCTGGATTGGTGTCTCAACTGCCCATACACTTACATGTCTTACTTCTCT
CAGGAATGTGACTGACACGGATATTCTAGCCCTGGAACGGCGACTGTGCAAACCATGGATA
TGATCATAAGCAAAAAGAAAAGGATGGCAATGGCACGGAGAACAATGTTCCAGAAGGGGAA
GTGCATAACAAACCATCAGGTTTCTGGGAATGATAAAAAGTGTACCATTCTCAGCATCAGG
AAGTGAAATCTTACTCTTATTTCAACAGGAAGTGGATGCTTTGGAAGAATTAAGCAGGCAGC
TTTTTCTGGAAACAGCTGATCTATATGCTACCAAGGAGAGAATAGAATACTCCAAAACCTTC
AAGGGGAAATATTTTAAATTTCTTGGTTACTTTTTCTCTATTTACTGTGTTTGGAAATTTT
CATGGCTACCATCAATATTGTTTTTGATCGAGTTGGGAAAACGGATCCTGTCAAGAGGCA
TTGAGATCACTGTGAATTATCTGGGAATCCAATTTGATGTGAAGTTTTGGTCCCAACACATT
TCCTTCATTCTTGTGGAATAATCATCGTCACATCCATCAGAGGATTGCTGATCACTCTTAC
CAAGTTCTTTTATGCCATCTCTAGCAGTAAGTCCTCCAATGTGCTGCTGCTATTAGCAC
AGATAATGGGCATGTACTTTGTCTCCTCTGTGCTGTGATCCGAATGAGTATGCCTTTAGAA
TACCGCACCATAACTACTGAAGTCCTTGGAGAACTGCAGTTCAACTTCTATCACCGTTGGTT
TGATGTGATCTTCTGGTCAGCGCTCTCTCTAGCATACTCTTCTCTATTGGCTCACAAAC
AGGCACCAGAGAAGCAAATGGCACCTTGAAACTTAAGCTACTACAGACTGTTAGAGGCCAGT
GGTTTCAAAATTTAGATATAAGAGGGGGGAAAAATGGAACCAGGGCCCTGACATTTTATAAAC
AAACAAAATGCTATGGTAGCATTTTTTACCTTCATAGCATACTCCTTCCCCGTGAGGTGATA
CTATGACCATGAGTAGCATCAGCCAGAACATGAGAGGGAGAATACTCAAGACAATACTCA
GCAGAGAGCATCCCGTGTGGATATGAGGCTGGTGTAGAGGCGGAGAGGACCAAGAACTAA
AGGTGAAAAATACACTGGAACCTCTGGGGCAAGACATGTCTATGGTAGCTGAGCCAAACACGT
AGGATTTCCGTTTTTAAAGTTTCATGGAAAAGTTTATAGCTTTGCCTTGAGATTGACTCATT
AAAATCAGAGACTGTAACAAAAAAAAAAAAAAAAAAAGGGCGGCCGCGACTCTAGAGTCG
ACCTGCAGAAGCTTGGCCCATGGCCCACTTGTTTATTGCAGCTTATAATG

FIGURE 173

MSFLIDSSIMITSQILFFGFGWLFFMRQLFKDYEIROQYVVQVIFSVTFAFSCTMFELIIFEI
LGVLNSSSRYPFHWMNLCVILLILVFMVPFYIGYFIVSNIRLLHKQRLLFSCLLWLTFMYFF
WKLGDPPFILSPKHGILSIEQLISRVGVIGVTTLMALLSGFGAVNCPYTYMSYFLRNVTDDI
LALERRLLQTMDMIISKKKRMAMARRTMFQKGEVHNKPSGFWGMIKSVTTSASGSENLTLIQ
QEVDALAEELSRQLFLETADLYATKERIEYSKTFKGKYFNFLGYFFSIYCVWKIFMATINIVF
DRVGKTDPTVRGIBITVNYLGIQFDVKFWSQHISFILVGIIIVTSIRGLLITLTKFFYAIS
SKSSNVIVLLAQIMGYFVSSVLLIRMSMPLEYRTIITEVLGELQFNFYHRWFDVIFLVSA
LSSILFLYLAHKQAPEKQMAP

Important features:

Signal peptide:

amino acids 1-23

Potential transmembrane domains:

amino acids 37-55, 81-102, 150-168, 288-311, 338-356, 375-398,
425-444

N-glycosylation sites.

amino acids 67-70, 180-183 and 243-246

Eukaryotic cobalamin-binding proteins

amino acids 151-160

FIGURE 174

CATGGGAAGTGGAGCCGGAGCCTTCCTTACACTCGCCATGAGTTTCCTCATCGACTCCAGCA
TCATGATTACCTCCCNGANACTATTTTTTGGATTGGGTGGCTTTTCTTCNGCGCCAATGTT
TAAAGACTATGAGATACGTCAGTATGTTGTACNGGTGATCTTCTCCGTGACGTTTGCCATTT
CTTGACCATGTTTGAGCTCATCATCTTTGAAATCTTNGGAGTATTGAATAGCAGCTCCCGT
TATTTTCACTGGAAAATGAACCTGTGTGTAATCTGCTGATCCTGGTTNTCATGGTGCCTTT
TTACATTGGCTATTTTATGTGAGCAATATCCGACTACTGCATAAACAACGACTGCTTTTTT
CCTGTCTCTTATGGCTGACCTTTATGTATTTCCAG

FIGURE 175

GTGTTGCCCTTGGGGAGGGGAAGGGGAGCCNGGCCCTTTCCTAAAAATTTGGCCAAGGGTTTC
TTTNTTGAATCCGGGTNNNGNATACCTTCCCAGAAAATATTTTTTGGATTTGGGGTAGNTT
TTTTTCATGCGCCAATTGTTTTAAAGACTATGAGATACGTCAGTATGTTGTACAGGTGATNTT
NTCCGTGACGTTTGCATTTTCTTGACCATGTTTGAGCTCATCATNTTTGAAATNTTAGGAG
TATTGAATAGCAGCTCCCGTTATTTTCACTGGAAAATGAACCTGTGTGTAATTCTGCTGATC
CTGGTTTTTCATGGTGCCTTTTTTACATTGGCTATTTTATTGTGAGCAATATCCGACTACTGCA
TAAACAACGACTGCTTTTTTCTGTCTNTTATGGCTGACCTTTATGTATTTNTTNTGGAAAN
TAGGAGATCCCTTCCCATTC

[illegible]

TTGCGCGAGGATCGTCCCCTGGCGCGGGCTCGGAGCCGCGACCTTTGGGGGGGCTCTCGGAGATTTCGTACTCTTTT
TGGCTCCCTCGTTCGTCGAACGTGCTCTTCTCAACGGGCTGTGCCCTTCAATCTGACAGTGTAGTGGTGCCTTGTGCGCA
GAGGGCGAGCGCAGGACGCTTCTTCGGCTTCTCTTGCGCCTGCACCGGCGATGTGAGCCCGACGCCAGGACGTG
GCTGTGCTGTGGGTGCTCCCGACGCTGCTGCTCTCTGGGACAGCGGCAATCGCATGTGAGGGCTTTCGCTT
CCGCTTGAGGCTGTGAGAGACTCAGTGTCTACAGATGGACATCGACAGGGAGCTGATATGCAAAAGGAGGACAA
GGAAACCACTGTTGTTGGGATTCAGTTCTCGGACGACGGGGCTGGGGGCAAGATTTGTTACTCTGTGCACACCGATA
TGAGGCAAGGCGACGAGTGTGACACGATCTGTGAGACGCGGGATATGATTGGTCTGCTGCTTTTGTGCTCAGGCAAGA
CCTGGCCACTCTCGGGATGAGTTTGATGGTGGGAATGGAAGTTTCTGTGAGGGACGCCCCCAAGGCCATGAACAATT
TGGGTTTCTGCGCAGAGGCAAGCTCGCGCTCTTCTCCCTGATAGGCCATCTCCTCTTTGGGGGCGCCAGAAAT
CTAATAATTGGAAGGGCAGCGCCAGGTTGAGCTCTGTGCAAGGCTCAGCGGACGTGGCAACCTGGACAGCGE
TCTCTACGAGCGCGGGGAGAGAGAGAGGACGAGCAACGCCGCTCATCCCGCTCTGCTCCAAAGACTATTTGGCT
TCCTATTGACTCGGGGAAAGGTTCTGTGCTGCAGAAAGACTGACGTTTGTGGTGTGAGCCCCCGGCCAACCA
CAAGGGTGTCTGTGGTTCATCTGTGCGCAAGGACAGCGCCAGTCGCTGTGTGCCAGAGTTATGCTGTCTGAGAGCGC
CTTGACCTCCGCTTTTGGTATCTGACTGGCTGTGGCTGACTTCAACGATGATGGCTGCGCAGACTGATAGTGGT
TGCCCTCCCTTTCTTTGAGTCCCAAGAGAGCTGGGGGGTCTGTGTATGTGTACTGAAACAGGGGGGCTACT
GGCTGGGATCTCCCTCTCGCGCTCTGTGGGCTCCCTGACTCATGTTCCGGATCAGCTGGCTGTCTCGGGGGA
CTTCAACCAAGATGGCTTTCCAGATATGCTAGTGGGCGCCCTTTGATGGTGTGGAAGATCTTCACTTACCA
TGGGACAGGCTCGGGGTTGTGCGCAAACTTCAACAGTGTCTGAGGAGCGGAGCTGTGGGCACTCAAGAGCTTCGG
TCTACTCCCTGTTCAGGACGCTTGATATGATGATGGGAACCAATACCTGACCTGTGCTGGTGGGCTCCTGGCTGCAC
CGCAGTGTCTTTCAGGCGCAGACCCATCTCATGACTCTCCATGAGTGTCTTATTGCTCCACGAGCATCGACCT
GGAGCGGCCCAATGCTGCTGGCGGCACTCGGTCTGTGTGTGACCTAAAGGTTCTGTTTTCAGTACATATTGACGTCCC
CAGCAGTATGCGCTACTGTGTGCTCTGCAATATGTGTATGATCGGCAACAGACCGGAGGCTCGGGGCGAGGT
TCCCGTGTGATGCTCTTCGAGCGTAACTGTGAAGAACCCCAAGCAGCAGGCTCTCGGGCAGCGGTGTGCTGTGAAGA
CCAGCATGACCGAGTCTGTGAGACGCGCATGTCTGACGTCCAGGAATAATCAAGACAGCTTCTCGGCCACTTGT
AGTGAACCTTGTCTTACAGTCTCCAGACCCCTCGGCTCGGCGACAGGCTCTCTGGCCAGGGGCTGTCTCCAGTGGC
CCCCACTCTCAATGGCCACACCGCCAGCAGCCAGCGGCGAGAGTCCACTTCTGAGGACAGGCTGTGGTGGAG
CAAGATCTCCCAAGCAATCTGACAGTGTGTCAACGCCGCTTCTGTACCCGGTTCAGCGACAGGAATCTCAACC
TCTGCCCTGTGATGTGTGATGGAAACACAGCCCTGTTTGACATGAGTGGGACAGCAATGATGGCTGGAGCTGAT
GGTCAACAACCTTGCCATCGAACCGACCGCCAGCGGACGTAGGGGATGGCCATGAGGACCCAGCTCTGGT
CATGCTCTGCTGACTCATGTGACTCTCAGGSGTCTGGGGCTGAGACCTGGGAGAGGCCATCTGCTCTGTCCAA
TGGAATGCTCTCCATCTGATGCTGAGCTGTGAGCTGGGAAACCCATGGAAGGCTGGCCAGTCACTTTCTACTCTAT
CCTTAGCACTCTCGGGATCAGCATGTGAGACACCGGAATCGAGGTAGAGCTGTGTTGGCCACGATCAGTGAGCA
GGAGCTGTGATCTGATCTCTGTCACAGCGCGTGTCTTCATGTAGCTGCACCTGTCCATTGTGAGAAATGGCCACTCC
CCAGCAACTCTCTTCTCTCTGTGTGTGGTGGGGGCGAGAGGACATCGAGCTGTGACGGGATGTGGGCGAGCAAGT
CACTATGAGTTACCGGTTTCCAAACAGGCGCTGCTCAGAACCTGGGCTCTGCTGCTCTCCCAACATCTG
GGCTCATGAGATTGCAAGTATGGGAATGGATGGTGTGCTGTATCCCATGCAAGTGTGAGTGTGGGCGCGGAGGCGCTGG
GCAGAAAGGGCTTTGCTCTCCAGCGCTCAACATCTCCAACTGTGATGTGACACATGAGGATGAGAGGCGCGCGGG
GCTGGAGGCGACCTCGACAGCAGGAGCTGTGTGAGCGGCGAGGAGCCAGCATGTCTGTGTGGCAGTGTCTCTGCT
TGAGAAGAAGAAAAAATCACTTCAGTGTGCGCCGGGGCAGCGCCAACTGTGTGGTGTCTCAGCTGCCACTCTA
CAGTCTTGAACCGCGCGGCTGTGTGCTCATCTCTGTGGGCGCTCTGTGGAACAGCACTTCTTGAGGAGTACTCAGT
TGTAAGTCCCTGGAAGTGTATTTCTGCGGCAACATCAGTGAAGTCTCTCAATAAGAACTTGATGCTCCGAGA
TGCCCTCCAGATGATCTCCAGTATGTTATCTTGAACCCCATGTGCTGTGTGCGAGGAAGGATGCCCTGTGTGGT
CATCTCTCTGGCTGTACTGGCTGGGCTGTGCTGTGTGCTGACTGCTGTGCTGTCTGCTGTGGAAGTGGGATTTCT
CAAAAGGGGGAAGCAAGCCGCGAGGCCACCGTGCCCACTGATCGGTGAGGATTTCTCGGGAAGACCGACAGCA
GTTCAAGGAGGAGAAGACGGGCAACATCTTGAGGAACAACATGGGGCAGCCCGCGGGGAGGGCCCGGATGAGCA
CCCATCTCTGGCTGTGACGGGCACTCCGAGTCTGGGCGCCGATGGGATCCGCGGACAGGCAAGCCGCTAGGTTCC
CATGCTCCAGGCTCGCTGTGGCTGCCCTCATTCCTTCCACAGATGTGCTCTGGGATGAAGAGGGTGTAGAGT
GGGCTGCTGGTGTGCATCAAGATTGTGGAGATGCGCTCTGACGGGCAACAGCTCTCCCAACAACAGAA
TCCTCCCAACCAACTTCCCTTAGAGTGTCTGTGAGATGAGATGGTAAATCAGGCAAGGGCCATGGGCTAGGG
TGAGAAGGGCAGGGGTGCTTCACTGACAAAGTGGGGGAGAGGGATCTCAATCCCTCTCTTCCATCTCACCTGT
GTAAACAGGACCCAGGACCTGCTCCCGGAGTGCCTTAACCTAGAGGCTCGGGAGAGGTTTGTCACTGATCA
CTCAGGCTGCTCTTCTCTAGTTTCTCCCTCTCATCTGACCTTAGTTTGTGTCATCAGTCTAGTGGTTCTGTGTGT
TTGCTCAATTATTAATAAATAATTTGGAGACAAAAAATAAATAAATAA

FIGURE 177

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA55737

><subunit 1 of 1, 1141 aa, 1 stop

><MW: 124671, pI: 5.82, NX(S/T): 5

MAGARSRDPWVGASGICYLFGSLLVELLFSRAVAFNLDVMGALRKEGEGPSLFGFSVALHRQL
QPRPQSWLLVGAPQALALPGQANRTGGLFACPLSLEETDCYRVDIDQGADMQEKESKENQWL
GVSVRSQGGPGGKIIVTCAHRYEARQQRVDQILETRDMIGRCFVLSQDLAIRDELDDGGGEWKFCEG
RPQGHEQFGFCQQGTAAAFSPDSHYLLFGAPGTYNWKGATARVELCAQGSADLAHLDDGPYEA
GGEKEQDPRLLIPVPANSYFGFSIDSGKGLVRAEELS FVAGAPRANHKGAVVILRKDSASRLV
PEVMLSGERLTSGFGYSLAVADLNSDGWPDLI VGAPYFFERQEELGGAVVYVLNQGHWAGI
SPLRLCGSPDSMFGISLAVLGDNLNQGDFDIAVGAPFDGDGKVF IYHGSSLGVVAKPSQVLE
GEAVGIKSFGYSLSGSLDMDGNQYPDLLVGSLADTAVLFRARPILHVSHEVSIAPRSIDLEQ
PNCAGGHSVCVDLRVCFSYIAVPSSYSPTVALDYVLDADTDRLRGQVPRVTFLSRNLEEPK
HQASGTVWLKHQHDRVCGDAMFQLQENVKDKLRAIVVTLSSYSLQTPRLRRQAPGQGLPPVAP
ILNAHQPSQTQRAEIHFLKQGCGEDKICQSNLQLVHARFCTRVSDTEFQPLPMDVDGTTALFA
LSGQPVIGLELMVTNLPSDPAQPDAGDDAHEAQLLVMLPDSLHYSGVRALDPAEKPLCLSN
ENASHVECELGNPMKRGAVTFYLLISTSGISIE TTELEVELLATISEQELHPVSARARVF
IELPLSIAGMAIPQQLFFSGVVGERAMQSERDVGSKVKYEVTVSNQGGQSLRTLGS AFLNIM
WPHEIANGKWLLPYPMQVELEGGQGPGQKGLCSPRPNILHLDVDSRDRRRRELEPPEQQEPGE
RQEPSMSWWPVSABKKKNITLDCARGTANCVVFSCLPLYSPDRAAVLVHVGRLWNSTFLEBY
SAVKSLVIVRANITVKSSIKNLMRLDASTVIPVMVYLDPMVAVVAEGVPPWWVILLAVLAGLL
VLALLVLLLWKMGFFKRAKHPEATVPQYHAVKIPREDRQQFKEEKTGITILRNWGS PRREGP
DAHPI LAADGHPGLGPDGHPGPGA

Important features:

Signal peptide:

amino acids 1-33

Transmembrane domain:

amino acids 1040-1062

N-glycosylation sites.

amino acids 86-89, 746-749, 949-952, 985-988 and 1005-1008

Integrins alpha chain proteins.

amino acids 1064-1071, 384-408, 1041-1071, 317-346, 443-465, 385-407, 215-224, 634-647, 85-99, 322-346, 470-479, 442-466, 379-408 and 1031-1047

FIGURE 178

CGCGCCGGGCGCAGGGAGCTGAGTGGACGGCTCGAGACGGCGGCGGTGCAGCAGCTCCAGA
AAGCAGCGAGTTGGCAGAGCAGGGCTGCATTTCCAGCAGGAGCTGCGAGCACAGTGTCTGGCT
CACAAACAAGATGCTCAAGGTGTCTAGCCGTAAGTGTGTGTGTGTGCAGCCGCTTGGTGCAGTCA
GTCTCTCGCAGCTGCCGCGGCGGTGGCTGCAGCCGGGGGCGGTTCGACGGCGGTAATTTTC
TGGATGATAAAACAATGGCTCACCACAATCTCTCAGTATGACAAGGAAGTCGGACAGTGGAAAC
AAATTCGAGACGAAGTAGAGGATGATTATTTCCGCACTTGGAGTCCAGGAAACCCCTTCGA
TCAGGCTTTAGATCCAGCTAAGGATCCATGCTTAAAGATGAAATGTAGTTCGCCATAAAGTAT
GCATTGCTCAAGATTCTCAGACTGCAGTCTGCATTAGTCACCGAGGCTTACACACAGGATG
AAAGAAGCAGGAGTAGACCATAGGCAGTGGAGGGGTCCCATATTATCCACCTGCAAGCAGTG
CCCAGTGGTCTATCCAGCCCTGTTTGTGGTTCAGATGGTCAATCTACTCTTTTCAGTGCA
AAGTAGAATATCAGGCATGTGTCTTAGGAAAAACAGATCTCAGTCAAAATGTGAAGGACATTGC
CCATGTCTTTCAGATAAGCCCAAGTACAAGCAGAAATGTTAAGAGAGCATGCAGTGACCT
GGAGTTTCAGGGAAGTGGCAAAACAGATTGCGGGACTGGTTCAAGGCCCTTCATGAAAGTGGAA
GTCAAAACAGAAGACAAAAACATTGCTGAGGCCTGAGAGAAGCAGATTTCGATACCAGCATC
TTGCCAATTTGCAAGGACTCACTTGGCTGGATGTTTAAACAGACTTGATACAAACTATGACCT
GCTATTGGACCAGTCAGAGCTCAGAAGCATTTACCTTGATAAGAATGAACAGTGTACCAAGG
CATTTCTTCAATTCCTGTGACACATACAAGGACAGTTTAATATCTAATAATGAGTGGTGTCTAC
TGCTTCCAGAGACAGCAAGACCCACCTTGCCAGACTGAGCTCAGCAATATTCAGAAGCGGCA
AGGGGTAAAGAAGCTCCTAGGACAGTATATCCCCCTGTGTGATGAAGATGGTTACTACAAGC
CAACACAATGTCTATGGCAGTGTGGACAGTGTCTGGTGTGTTGACAGATATGGAAATGAAGTC
ATGGGATCCAGAATAAATGGTGTTCAGATTGTGCTATAGATTTTGAGATCTCCGAGATTT
TGCTAGTGGCGATTTTCATGAATGGACTGATGATGAGGATGATGAAGACGATATTATGAATG
ATGAAGATGAAATGAAGATGATGATGAAGATGAAGGGATGATGATGATGGTGGTGTGAC
CATGATGTATACATTTGATTGATGACAGTTGAAATCAATAAATTTCTACATTTCTAATATTTA
CAAAAATGATAGCCTATTTAAATTTATCTTCTCCCCAATAACAAAATGATTCTAAACCTCA
CATATATTTTGTATAATTATTTGAAAAATTCAGCTAAAGTTATAGAACTTTATGTTTTAAAT
AAGATCATTGTCTTTGAGTTTTTATATTCTTACACAAAAAGAAAATACATATGCAGTCTA
GTCAGACAAAATAAAGTTTTGAAGTGCTACTATAATAAATTTTTTACGAGAACAACTTTGT
AAATCTTCCATAAGCAAAATGACAGCTAGTGTCTGGGATCGTACATGTTAATTTTTTTGAAG
ATAATTTCAAGTGAATTTAAATAAATAAATTTTTAATGACCTGGGTCTTAAGGATTTAGG
AAAAATATGCATGCTTTAATGCAATTTCAAAGTAGCATCTTGCTAGACCTAGATGAGTCAG
GATAACAGAGAGATACCACATGACTCCAAAAAAAAAAAAAAAAA

FIGURE 179

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA49829

><subunit 1 of 1, 436 aa, 1 stop

><MW: 49429, pI: 4.80, NX(S/T): 0

MLKVSAVLCVCAAAWCSQSLAAAAVAAGGRSDGGNFLDDKQWLTTISQYDKEVGQWNKFR
DEVEDDYFRTWSPGKPFQALDPAKDPCCLKMKCSRHKVCIAQDSQTAVCISHRRLTHRMKEA
GVDHRQWRGPILSTCKQCPVVYPSPVCGSDGHTYSFQCKLEYQACVLGKQISVKCEGHCP
SDKPTSTSRNVKRACSDLEFREVANRLRDWFKALHESGSQNKKTILLRPERSRFDTSILPI
CKDSLGMFNRDLDTNYDLLDQSELRSIYLDKNEQCTKAFFNSCDTYKDSLISNNEWCYCFQ
RQQDPPCQTELSNIQKRQGVKKLLGQYIPLCEDGYYKPTQCHGSGVQCWCVDTRYGNEVMGS
RINGVADCAIDFEISGDFASGDFHEWTDDEDEDDIMNDEDEIEDDEDEBDDDDGGDDHVDYI

Important features:

Signal peptide:

amino acids 1-16

Leucine zipper pattern.

amino acids 246-267

N-myristoylation sites.

amino acids 357-362, 371-376 and 376-381

Thyroglobulin type-1 repeat proteins

amino acids 353-365 and 339-352

Figure 1 consists of seven sub-graphs labeled (a) through (g), each showing a time course of a different physiological parameter over a 10-minute period. The y-axis for all graphs ranges from 0 to 100. The x-axis for all graphs ranges from 0 to 10 minutes. The graphs show that HR, SV, CO, and MAP increase during the intervention, while PVR, SVR, and the PVR/SVR ratio decrease.

Time (min)	(a) HR (b/min)	(b) SV (ml)	(c) CO (l/min)	(d) MAP (mmHg)	(e) PVR (mmHg)	(f) SVR (mmHg)	(g) PVR/SVR ratio
0	70	50	3.5	90	100	100	1.0
1	70	50	3.5	90	100	100	1.0
2	70	50	3.5	90	100	100	1.0
3	70	50	3.5	90	100	100	1.0
4	70	50	3.5	90	100	100	1.0
5	70	50	3.5	90	100	100	1.0
6	70	50	3.5	90	100	100	1.0
7	70	50	3.5	90	100	100	1.0
8	70	50	3.5	90	100	100	1.0
9	70	50	3.5	90	100	100	1.0
10	70	50	3.5	90	100	100	1.0

[illegible]

FIGURE 181

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA52196

><subunit 1 of 1, 229 aa, 1 stop

><MW: 26017, pI: 4.73, NX(S/T): 0

MGDKIWLPEFPVLLLAALPPVLLPGAAGFTPSLSDSDFTFITLPAGQKECFYQPMPLKASLEIEY
QVLDGAGLDIDFHLASPEGKTLVFEQRKSDGVHTVETEVGDYMFCDNTFTSTISEKVIFTEL
ILDNMGEQAQEDEDWKYITGTDILDMKLEDILESINSIKSRLSKSGHIQILLRAFEARDRN
IQESNFDRVNFWSMVNLVVMVVVSAIQVYMLKSLFEDKRKSRT

Important features:

Signal peptide:

amino acids 1-23

Transmembrane domain:

amino acids 195-217

N-myristoylation site.

amino acids 43-48

Tyrosine kinase phosphorylation site.

amino acids 55-62

FIGURE 182

CCATCCCTGAGATCTTTTATAAAAAACCCAGTCTTTGCTGACCAGACAAAGCATACCAGAT
CTCACCAGAGAGTGCAGACACTATGCTGCCTCCCATGGCCCTGCCAGTGTGTCTGGATG
CTGCTTTTCTGCCTCATTTCTCTGTGTGAGGTTCAAGGTGAAGAAAACCCAGAAGGAACTGCC
CTCTCCACGGATCAGCTGTCCCAAAGGCTCCAAGGCCTATGGCTCCCCCTGCTATGCCTTGT
TTTTGTCAACAAAATCCTGGATGGATGCAGATCTGGCTTGCCAGAAGCGGCCCTCTGGAAAA
CTGGTGTCTGTGCTCAGTGGGGCTGAGGGATCCTTCGTGTCCTCCCTGGTGAGGAGCATTAG
TAACAGCTACTCATACATCTGGATTGGGCTCCATGACCCACACAGGGCTCTGAGCCTGATG
GAGATGGATGGGAGTGGAGTAGCACTGATGTGATGAATTACTTTGCATGGGAGAAAAATCCC
TCCACCATCTTAAACCCTGGCCACTGTGGGAGCCTGTCAAGAAGCACAGGATTTCTGAAGTG
GAAAGATTATAACTGTGATGCAAAGTTACCCATATGTCTGCAAGTTCAGGACTAGGGCAGGT
GGGAAGTCAGCAGCCTCAGCTTGGCGTGCAGCTCATCATGGACATGAGACCAGTGTGAAGAC
TCACCCCTGGAAGAGAATATTTCTCCCAAACCTGCCCTACCTGACTACCTTGTGATGATCCTCC
TTCTTTTTCTTTTCTTACCTTCATTTGAGGCTTTTCTGTCTTCCATGTCTTGAGATC
TCAGAGAATAATAAAAAATGTTACTTTATAAAAAAAAAAAAAAAAAAAAAA

1000 900 800 700 600 500 400 300 200 100 0

<subunit 1 of 1, 175 aa, 1 stop

MLPPMALPSVSWMLLSCLILLCQVQGEETQKELPSPRISCPKGSKAYGSPCYALFLSPKSWM

TDVMNYFAWEKNPSTILNPGHCGSLRSTGFLKWKDYNCDAPYVCKFKD

Signal peptide:

C-type lectin domain signature.

amino acids 146-171

FIGURE 184

CCAGTCTGTGCGCCACCTCACTTGGTGTCTGCTGTCCCCGCCAGGCAAGCCTGGGGTGAGAGC
ACAGAGGAGTGGGCCGGGACCATGCGGGGGACGCGGCTGGCGCTCCTGGCGCTGGTGCTGGC
TGCCTGCGGAGAGCTGGCGCCGGCCCTGCGCTGCTACGTCTGTCCGGAGCCCACAGGAGTGT
CGGACTGTGTCACCATCGCCACCTGCACCACCAACGAAACCATGTGCAAGACCACACTCTAC
TCCCGGGAGATAGTGTAACCCCTTCCAGGGGGACTCCACGGTGACCAAGTCCTGTGCCAGCAA
GTGTAAGCCCTCGGATGTGGATGGCATCGGCCAGACCCCTGCCCGTGTCTGTGCAATACTG
AGCTGTGCAATGTAGACGGGGCGCCCGCTCTGAACAGCCTCCACTGCGGGGCCCTCACGCTC
CTCCCACTCTTGAGCCTCCGACTGTAGAGTCCCCGCCACCCCCATGGCCCTATGCGGCCCA
GCCCCGAATGCCTTGAAGAAGTGCCCCCTGCACCAGGAAAAAAAAAAAAAAAAA

0992403-101601
109101-20432650

FIGURE 185

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56405

<subunit 1 of 1, 125 aa, 1 stop

<MW: 13115, pI: 5.90, NX(S/T): 1

MRGTRLALLALVLAACGELAPALRCYVCPEPTGVSDCVTIATCTTNETMCKTTLYSREIVYP
FQGDSTVTKSCASKCKPSDVGIGQTLPVSCNTELCNVDGAPALNSLHCGALTLLPLLRL

Important features:

Signal peptide:

amino acids 1-17

N-glycosylation site.

amino acids 46-49

FIGURE 186

CTGCAGTCAGGACTCTGGGACCGCAGGGGGCTCCCGGACCCTGACTCTGCAGCCGAACCGGC
ACGGTTTCGTGGGGACCCAGGCTTGCAAAGTGACGGTCATTTTCTCTTTCTCCCTCTT
GAGTCCCTCTGACATGATGGCTCTGGGCGCAGCGGGAGCTACCCGGGTCTTTGTCCGGATGG
TAGCGGCGGCTCTCGGCGGCCACCTCTGCTGGGAGTGAGCGCCACCTTGAACCTCGGTTCTC
AATTCCAACGCTATCAAGAACCTGCCCCACCGCTGGGCGGCGCTGCGGGGCACCCAGGCTC
TGCAGTCAGCGCCGCGCGGGAATCCTGTACCCGGGCGGGAATAAGTACCAGACCATTGACA
ACTACCAGCCGTACCCGTGCGCAGAGGACGAGGAGTGCGGCACTGATGAGTACTGCGCTAGT
CCCACCCGCGGAGGGGACGCAGGCGTGCAAATCTGTCTCGCCTGCAGGAAGCGCCGAAAACG
CTGCATGCGTCACGCTATGTGCTGCCCCGGAATTACTGCAAAAATGGAATATGTGTGTCTT
CTGATCAAAAATCATTTCCGAGGAGAAATTGAGGAAACCATCACTGAAAGCTTTGGTAATGAT
CATAGCACCTTGGATGGGTATTCAGAAGAACCACCTTGTCTTCAAAAATGTATCACACCAA
AGGACAAGAAGGTTCTGTTTGTCTCCGGTCATCAGACTGTGCCTCAGGATTGTGTTGTGCTA
GACACTTCTGGTCCAAGATCTGTAAACCTGTCTGAAAGAAGGTCAAGTGTGTACCAAGCAT
AGGAGAAAAGGCTCTCATGGACTAGAAATATTCAGCGTTGTTACTGTGGAGAAGGTCTGTC
TTGCCGGATACAGAAAGATCACCATCAAGCCAGTAATTCTTCTAGGCTTCACACTTGTGAGA
GACACTAAACCAGCTATCCAAATGCAGTGAACCTCTTTTATATAATAGATGCTATGAAAACC
TTTTATGACCTTCATCAACTCAATCCTAAGGATATACAAGTTCTGTGGTTTCAGTTAAGCAT
TCCAATAACACCTTCCAAAAACCTGGAGTGAAGAGCTTTGTTTCTTTATGGAACCTCCCCTG
TGATTGCAGTAAATTACTGTATTGTAAATTCTCAGTGTGGCACTTACCTGTAAATGCAATGA
AACTTTTAATTATTTTCTAAAGGTGCTGCACTGCCTATTTTTCCTCTTGTATGTAAATTT
TTGTACACATTGATTTGTATCTTGACTGACAAATATTCTATATTGAACTGAAGTAAATCATT
TCAGCTTATAGTTCTTAAAAGCATAACCTTTACCCCATTTAATTCTAGAGTCTAGAACGCA
AGGATCTCTTGGAATGACAAATGATAGGTACCTAAAAATGTAACATGAAAATACTAGCTTATT
TTCTGAAATGTACTATCTTAATGCTTAAATTATATTCCCTTTAGGCTGTGATAGTTTTTGA
AATAAAATTTAACATTTAAAAA

FIGURE 187

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA57530

<subunit 1 of 1, 266 aa, 1 stop

<MW: 28672, pI: 8.85, NX(S/T): 1

MMALGAAGATRVFVAMVAAALGGHPLLGVSATLNSVLNSNAIKNLPPLGGAAGHPGSAVSA
APGILYPGGNKYQTIDNYQPYPCAEDEECGTDEYCASPTRGGDAGVQICLACRKRRCMRH
AMCCPGNYCKNGICVSSDQNHFRGEIEETITESFGNDHSTLDGYSRRITLSSKMYHTKGQEG
SVCLRSSDCASGLCCARHFWSKICKPVLKEGQVCTKHRRKGSHGLEIFQRCYCGEGLSCRIQ
KDHHQASNSSRLHTCQRH

Important features:

Signal peptide:

amino acids 1-23

N-glycosylation site.

amino acids 256-259

Fungal Zn(2)-Cys(6) binuclear cluster domain

amino acids 110-126

FIGURE 188

TGTGTTTCCCTGCAGTCAGAATTTGGGACNGCAGGGGTTCCCGGACCTGATTTTGCAGCGGA
ACGGGAAGGTTTTGTGGGACCCAGGTTGAAATGACGGTCATTTTTTTTCTTCTCCTTCNG
GAGTCCTTNTGAGANGATGGTTTTTGGGCGCAGCGGGAGCTAACCCGGTTTTTTGTNGCGATG
GTAGCGGCGGTTTTTCGGCGGCCACCTTNTGCTGGGAGTGAGCGCCACCTTGAATCGGTTTTTC
AATTCCAACGNTATCAAGAACCTGCCCCCACCNTGGGCGGCGCTGCGGGGCACCCAGGNTT
TGCAGTCAGCGCCGCGCCGGAATCCTGTACCCGGGCGGGAATAAGTACCAGACCATTGACA
ATTACCAGCCGTACCCGTGCGCAGAGGACGAGGAGTGCGGCACTGATGAGTACTGCGCTAGT
CCCACCCGCGGAGGGGANGCGGGCGTGCAAATNTGTNTNGCCTGCAGGAAGCGCCGAAAACG
CTGCATGCGTCANGCTATGTGCTGCCCCGGAATTACTGCAAAAATGGAATATGTGTGTTTT
CTGATCAAAATCATTTCCGAGGAGAAATTGAGGAAACCATCACTGAAAGCTTTGGTAATGAT
CATAGCACCTTGGATGGG

FIGURE 189

GAGGAACCTACCGGTACCGGCCGCGCTGGTAGTCGCGCGGTGGCTGCACCTCACCAATCCCCTGCGCCGCGG
CTGGGCGCTCGGAGAGTGCCTGTGCTTCTCTCCTGCACGCGGTGCTTGGGCTCGGCCAGGCGGGGTCCCGCGCCA
GGGTTTGAGGATGGGGAGTAGCTACAGGAAGCGACCCCGCGATGGCAAGGTATATTTTGTGGAATGAAAGGA
AGTATTAGAAATGAGCTGAAGACCATTTACAGATTAAATATTTTGGGGACAGATTTGTGATGCTTTGATTCACCT
TGAAGTAATGTAGACAGAACTTCTCAAAATTTGCATATTACATCAACTGGAACCGCAGTGAATCTTAATGTTTAC
TTAAATCAGAACTTGCAATAGAAAGAGATGGGAGTCTGGTTAAATAAAGATGACTATATCAGAGACTTTGAAAG
GATCAATCTCTGTTTTCTGTAGTGTATATGGCCATTTTAGTGGGACAGATCAGGATTTTACAGTTTACTTGG
AGTGTCCAAAACCTCAGCAGTAGAGAAATAAGACAAGCTTTCAAGAAATTTGGCAATTTGAAATTTACATCTGTATA
AAAGCCGAATAACCCAAATGCATCGCGGATTTTAAAAATAAATAGAGCATATGAAGTACTCAAGATGAAGA
TCTACGGAAAAAGATGACAAATATGGAGAAAGGGAAGTCTGAGGATTAATCAAGGTGGCCAGTATGAAAGCTGGAA
CTATTATGCTTATGATTTTGGTATTTTATGATGATGATCCTGAAATCATAACATTGGAAGAAAGAGAATTTGATGC
TGCTGTTAAATCTGGAGAAGCTTGGTTTGTAAATTTTACTCCCGAGCTGTTCCACACTGCCATGATTTAGCTCC
CACATGGAGAGACTTTGCTAAAGAAAGTGGATGGGTTACTTCGAATTTGGAGTGGCTTAACTGCTGGTGATGATAGAAT
GCTTTGGCGAATGAAAGGAGTCAACAGCTATCCAGTCTCTTCATTTTTCGGTCTGGAATGGCCCAGTGAATA
TCATGGAGACAGATCAAGAGGAGATTTTAGTGAGTTTTCGAATGCAGCATGTTAGAAGTACAGTGACAGAACTTTG
GACAGGAAATTTTGTCAACTGCCATCAAAAGCTGCTTTTGTCTGCTGATATTTGGCTGGCTGATCACTTTTGTTCAAA
AGGAGGAGATGTTTGTGACTTCACAGACAGCATCAGGCTTAGTGGCATGTTGTTTCTCAACTCATTTGGATGCTAA
AGAAATATATTTGGAAGTATATCATATCTTCCAGATTTTGAACACTTTCCGCGAAACACACTAGAGGATCGTTT
GGCTCATCATCGGTGGCTGTTATTTTTCATTTTGGAAAAAATGAAATTCAAATGATCTGAGCTGAAAAAAT
AAAAACTCTACTAAAAATCATCATATTTCAAGTTGGCAGGTTTGTACTGTTCTCTGCACAGACATCTGTAGTAA
TCTGTATGTTTTTTCAGCCGCTTCAGCATATTTAAAGGACAAGGAAACAAAGATATGAAATTCATCATGGAAA
GAAGATCTTATGATATACCTTGCTTTTGCCAAAGAAAGTGTGAATTTCTCATGTTACACGCTTTGGACCTCAAAA
TTTTCTGCGCAATGACAAAGAACCATGGCTTGTGATTTCTTGCCCCCTGGTATGCCACATGCGAGCTTTACT
ACCAGAGTTACGAAGAGCATCAATCTTCTTATGTCAGCTTAAGTTTGGTACACTAGATTTGACAGTTCATGTA
GGGACTCTGTAACTGTATAAATCTCAGGCTTATCCAAACACAGTGGTATTCAACCACTCCACATCTCATGATGA
TGAAGGACATCACTCTGCTGCAACAAATCTTGGAGTTTCAAGAGGATCTTATGAATCCTTCAGTGGTCTCCCTTAC
ACCCACCACTTCAACGAACATAGTTACACAAAGAAAAACACACAGGAAGTCTGGATGTTTGAATTTCTATTCTCCGTG
GTGCTCATCTTGGCAGCTCTTAAATGCAAGAAATGGAAAGAAATGGCCCGGACATTACCTGGAGCTATCAACGTGGG
CAGTATAGATTGCCAAAGTATCATTTCTTTTGTGCCCAGGAAAAAGCTTCAAAGATACCCGTGAGATGAATTTT
TCCCCCAAAATCAAATAAAGCTTATCAGTATCAGGTTACAAATGGTTGGAATAGGGATGCTTTTCCCTGAGAAT
CTGGGCTCAGGAATTTTACTCTCAAGTATCCACAGATCTAACACCTCAGACTTTTCAGTGAAAAAGTTTACAAAG
GAAAAATCATTTGGGTGATTGATTTCTATGCTCCTTGGTGTGGACCTTGCCAGAATTTTGCTCCAGAATTTTGAGCT
CTTGGCTAGGATGATTAAAGGAAAGTGAAAGCTGGAAAAAGTAGAGTCTCAGGCTTATGCTCAGACATGCCAGAA
AGCTGGGATCAGGGCCTATCCAAGTGTAAAGTTTATTTCTACGAAAGAGCAAGAAAGAAATTTCAAGAAAGAGCA
GATAAATACAGAGATGCAAAAGCAATCGCTGCCCTTAAAGTGAAGAAATTTGGAAGAACTCTCCGAAATCAAGGCCAA
GAGGAATAAGGATGAACTTTGAATATGTTGAAGATGAAGAAAAAGTTTAAAGAAATTTCTGACAGATGACATCAG
AAGACACCTATTTAGAATGTTACTTTATGATGGGAATGAATGAACATTTTACTGCTGACGTTGTACTGCCA
GAAATTTCTACAGCAGCTGGTGTAAAAGAGGGTCTGCAAACTTTTCTGTAAGGGGCGGTTTATAAATATTTTA
GACTTTGCGAGCTATAATATATGTTTCCACATAGAGAACAGAAATAGAGTATCATGATATCTTTGTTATTTGCT
TTTTAACAACTTTAAAAAATATTAAGACGATTTCTAGCTCAGAGCCATACAAAGTAGGGCTGGATTCAGTCCATG
GACCATAGATTGCTGCTCCCTCGACGGACTTATAATGTTTCAAGTGGCTGGCTTGAACATGAGTCTGCTGTGCT
ATCTACATAAATGCTCAAGTTGATATAAAGTCCACTTCCCTTCAAGTTTGTGGCTGACCTGAAAGAGGTAAT
TAGTTTTTGGTCACTGTTCTTCTCAAAATGCTATCCCTAACCATATTTATATTCGTTTTTAAAAACCCCAT
GATGTGGCAGAGTAAACAAACCTGTTATGCTGTATTAATATGAGGAGATTTCTCATTTGTTTCTTCTTCA
AAGGTTGAAAAAATGCTTTTAAATTTTCACTCAGCCGAGAAACAGTGCAGCAGTATATGTGCACACAGTAAGTACAC
AAATTTGAGCAACAGTAAAGTGCACAAATCTCTGAGTTTGTCTGTATCACTCAGGAAAAACCTGAGGGGAAAAAATTA
TAGCAATTAAGTGGGCTGTGAGAGTATCTCAAAATATGTTATCAAGTATTTAGAGTTCTATATTTTAAAGATATA
TGTTGTTCAATGATTTTCTGAAATGCTTTTCATGAAATTTTCCCACTGATAGTTGATTTTGGAGGACATCAATAT
TTACATATTTGCCTTTGTGAACTTTGTTTGGACCTGTATCTTTATTTACATTTGCTGTTTCTTTTCAAGTTTGG
TTTTTCACTCCTGCTCCAGTCTATTTATTTATTTCAAAATAGGAAAAATTAATTTTACAGGTTGTTTTTACTGTAGCTTAT
AATGATATCTGAGTTATTTCAAGTTTACTGTGTCAGAGGGCTGCCCTTTTCAGATAAATATGACATAATA
ACTGAAGTTATTTTATAAGAAATCAAGTATATAAATCTAGGAAGGGATCTTCTGATTTCTGTGTTGTTTGA
CTCAAGAAATCAGAAATTTTGTGAGTAACTGATGTTGTTTATGATTAATTTCAAGTGTACAGAAATGGTAAAAAT
CCAATCAGTCAAAAGAGGTCAATGAATTAAGAGGCTTGCAACTTTTCAAAAAA

FIGURE 190

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56439

<subunit 1 of 1, 747 aa, 1 stop

<MW: 86127, pI: 7.46, NX(S/T): 2

MGVWLNKDDYIRD LKRIILCFLIVYMAILVGT DQDFYSLLGVSKTASSREIRQAFKKLALKL
HPDKNPNNPNNAHGDFL KINRAYEVLKDEDLRKKYDKYGEKLEDNQGQYESWNYRYDFGI
YDDDEPIITLERREFDAAVNSGELWVFVNFYSPGCSHCHDLAPTWRDFAKEVDGLLRIGAVNC
GDDRMLCRMKGVNSYPSLFIFRSGMAPVKYHGDRSKESLVSFAMQHVRSVTTELWTGNFVNS
IQTAFAAGIGWLITFC SKGGDCLTSQTRLRLSGMLFNLSDAKEIYLEVIHNLPDFELLSAN
TLEDRLAHHRWLLFFHFGKNENSNDPELKKLKTLLKNDHIQVGRFDCSSAPDICSNLYVFQ P
SLAVFKGQGTKEYEIHGKKILYDILAFAKESVNSHVTTLG PQNFNPANDKEPWLVDFFAPWC
PPCRALLPELRRASNLLYGQLKFGTLDCTVHEGLCNMYNIQAYPTTVVFNQSNIEHEGHH S
AEQILEFIEDLMNPVS VSLTPTTFNELVTQRKHNEVWMVDFYSPWCHPCQVLMPEWKRMART
LTGLINVGSIDCQYH SFCAQENVQRYPEIRFFPPKSNKAYQYHSYNGWNRDAYSLRIWGLG
FLPQVSTDLTPQT FSEKVLQGNHWWIDFYAPWCGPCQNFAP EPELLARMIKGVKAGKVDC
QAYAQTCQKAGIRAYPTVKFYFYERAKRNFQEEQINTRDAKAI AALISEKLETLRNQGRNKDEL

Important features:

Endoplasmic reticulum targeting sequence.

amino acids 744-747

Cytochrome c family heme-binding site signature.

amino acids 158-163

Nt-dnaJ domain signature.

amino acids 77-96

N-glycosylation site.

amino acids 484-487

FIGURE 191

AGACAGTACCTCCTCCCTAGGACTACACAAGGACTGAACCAGAAGGAAGAGGACAGAGCAAA
GCCATGAAACATCATCCTAGAAATCCTTCTGCTTCTGATCACCATCATCTACTCCTACTTGGA
GTCGTTGGTGAAGTTTTTTCATTCTCAGAGGAGAAAACTGTGGCTGGGGAGATTGTTCTCA
TFACTGGAGCTGGGCATGGAATAGGCAGGCAGACTACTTATGAATTTGCAAAACGACAGAGC
ATATTGTTCTGTGGGATATTAATAAGCGCGGTGTGGAGGAACTGCAGCTGAGTGCCGAAA
ACTAGGCGTCACTGCGCATGCGTATGTGGTAGACTGCAGCAACAGAGAAGAGATCTATCGCT
CTCTAAATCAGGTGAAGAAAGAAGTGGGTGATGTAACAATCGTGGTGAATAATGCTGGGACA
GTATATCCAGCCGATCTTCTCAGCACCAAGGATGAAGAGATTACCAAGACATTTGAGGTCAA
CATCCTAGGACATTTTTGGATCACAAAAGCACTTCTTCCATCGATGATGGAGAGAAATCATG
GCCACATCGTCACAGTGGCTTCAGTGTGCGGCCACGAAGGGATTCTTACCTCATCCCATAT
TGTTCCAGCAAATTTGCCGCTGTGGCTTTCACAGAGGTCTGACATCAGAACTTCAGGCCTT
GGGAAAACTGGTATCAAAACCTCATGTCTCTGCCAGTTTTTGTGAATACTGGGTTACCA
AAAATCCAAGCACAGAATATATGGCTGTATTGGAGACAGATGAAGTCGTAAGAAGTCTGATA
GATGGAATACTTACCAATAAGAAAATGATTTTTGTTCATCGTATATCAATATCTTTCTGAG
ACTACAGAAGTTTCTTCTGAACGCGCCTCAGCGATTTTAAATCGTATGCAGAATATTCAAT
TTGAAGCAGTGGTTGGCCACAAAATCAAATGAAATGAATAAAATAAGCTCCAGCCAGAGATG
TATGCATGATAATGATATGAATAGTTTTCGAATCAATGCTGCAAAGCTTATTTACATTTTT
TCAGTCTTGATAATATTA AAAACATTGGTTTGGCACTAGCAGCAGTCAAACGAACAAGATTA
ATTACCTGTCTTCTGTCTCAAGAATATTTACGTAGTTTTTTCATAGGTCGTTTTTCTCTT
TCATGCCTCTTAAAACTTCTGTGCTTACATAAACATACTTAAAAGGTTTTCTTTAAGATAT
TTATTTTTTCATTTAAAGGTGGACAAAAGCTACCTCCCTAAAAGTAAATACAAAGAGAACT
TATTTACACAGGGAAGGTTTAAGACTGTTCAAGTAGCATTCCAATCTGTAGCCATGCCACAG
AATATCAACAAGAACACAGAATGAGTGCACAGCTAAGAGATCAAGTTTCAGCAGGCAGCTTT
ATCTCAACCTGGACATATTTTAAGATTTCAGCATTTGAAAGATTTCCCTAGCCTCTTCTTTTT
TCATTAGCCCAAAACGGTGCAACTCTATTCTGGACTTTATTACTTGATTCTGTCTTCTGTAT
AACTCTGAAGTCCACCAAAAGTGGACCCCTCTATATTCTCTCCCTTTTTTATAGTCTTATAAGA
TACATTATGAAGGTGACCGACTCTATTTTAAATCTCAGAATTTTAAAGTTCTAGCCCCATGA
TAACCTTTTTCTTTGTAATTTATGCTTTCATATATCCTTGGTCCCAGAGATGTTTAGACAAT
TTTAGGCTCAAAAATTAAGCTAACACAGGAAAAGGAAGTGTACTGGCTATTACATAAGAAA
CAATGGACCCAAGAGAAGAA

MNIIILEILLLLITIIYSYLESIVKFFIPQRRKSVAGEIVLITGAGHGIGRQTTYFEAKRQSI
 LVLWDINKRGVEETAABCRKLGVTAAHAYVVDSCNREEIYRSLNQVKKEVGDTVIVVNNAGTV
 YPADLLSTKDEEITKTFEVNILGHFWITKALLPSMMERNHGHIVTVASVCGHEGIPYLPYPC
 SSKFAAVGFHRLTSELQALGKTGIKTSCPCPVFNTGFTKNPSTRLPVPLETDEVVRSRID
 GILTNNKMI FVPSYINIFRLQKFLPERASAILNRMONIOFEAVVGHKIKMK

Signal peptide:

amino acids 1-19

62

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 30-33 and 58-61

Short-chain alcohol dehydrogenase family protein

amino acids 165-202, 37-49, 112-122 and 210-219

FIGURE 193

CGGCGGCGGCTGCGGGCGCGAGGTGAGGGGCGCGAGGTGAGGGGCGCGAGGTTCCAGCAGG
ATGCCCCGGCTCTGCAGGAAGCTGAAGTGAGAGGCCCGGAGAGGGCCAGCCCGCCGGGGC
AGGATGACCAAGGCCCGGCTGTTCCGGCTGTGGCTGGTGCTGGGGTCCGTGTTTCATGATCCT
GCTGATCATCGTGTACTGGGACAGCGCAGGCGCCGCGCACTTCTACTTGCACACGTCTCTCT
CTAGGCCGCGACACGGGGCCGCCGCTGCCCACGCCCGGGCCGGACAGGGACAGGGAGCTCACG
GCCGACTCCGATGTGACAGGTTTCTGGACAAGTTTCTCAGTGCTGGCGTGAAGCAGAGCGA
CCTTCCCAGAAAGGAGACGGAGCAGCCGCTGCGCCGGGGAGCATGGAGGAGAGCGTGAGAG
GCTACGACTGGTCCCCGCGCGACGCCCGGCGCAGCCAGACCAGGGCCGGCAGCAGGCGGAG
CGGAGGAGCGTGCTGCGGGGCTTCTGCGCCAACCTCCAGCCTGGCCTTCCCCACCAAGGAGCG
CGCATTTCGACGACATCCCCAACTCGGAGCTGAGCCACCTGATCGTGGACGACCCGGCACGGGG
CCATCTACTGCTACGTGCCCAAGGTGGCCTGCACCAACTGGAAGCGCGTGATGATCGTGCTG
AGCGGAAGCCTGTGCAACGCGGTGCGCCCTACGCGACCCGCTGCGCATCCCGCGCGAGCA
CGTGACACAACGCCAGCGCGCACCTGACCTTCAACAAGTTCTGGCGCGCTACGGGAAGCTCT
CCCGCCACCTCATGAAGGTCAAGCTCAAGAAGTACACCAAGTTCTCTTCTGTCGCGACCCC
TTCGTGCGCCTGATCTCGCCTTCCGCAGCAAGTTCGAGCTGGAGAACGAGGAGTTCTACCG
CAAGTTTCGCCGTGCCCATGTGCGGCTGTACGCCAACACACCAGCCTGCCCGCCTCGGCGC
GCGAGGCCTTCCGCGCTGGCCTCAAGGTGTCTTTCGCCAATTCATCCAGTACCTGTGGAC
CCGCACACGGAGAAGCTGGCGCCCTTCAACGAGCACTGGCGGCAGGTGTACCGCCTCTGCCA
CCCGTCCCAGATCGACTACGACTTCGTGGGGAAGCTGGAGACTCTGGACGAGGACGCCGCGC
AGCTGCTGCAGCTACTCCAGGTGGACCGGCAGCTCCGCTTCCCCCGAGCTACCGGAACAGG
ACCGCCAGCAGCTGGGAGGAGGACTGGTTCGCCAAGATCCCCCTGGCCTGGAGGCAGCAGCT
GTATAAACTCTACGAGGCCGACTTTGTCTCTTCGGCTACCCCAAGCCCAGAAACCTCTCC
GAGACTGAAGGCTTTCGCGTTGCTTTTTCTCGCGTGCCTGGAACTGACGCACGCGCACTCC
AGTTTTTTTATGACCTACGATTTTGCAATCTGGGCTTCTTGTTCACTCCACTGCCTCTATCC
ATTGAGTACTGTATCGATATTGTTTTTTAAGATTAATATATTTTCAGGTATTTAATACGA

FIGURE 194

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56112

<subunit 1 of 1, 414 aa, 1 stop

<MW: 48414, pI: 9.54, NX(S/T): 4

MTKARLFRLWLVLGSVFMILLIIIVWDSAGAAHFYLHTSFSRPHTGPPLPTPGPDRDRELT
DSDVDEFLDKFLSAGVKQSDLPRKETEQPPAPGSMEEESVRGYDWSPRDARRSPDQGRQQAER
RSVLRGFCANSSLAFPTKERAFDDIPNSELSHLIVDDRHGAIYCYVPKVACTNWKRMIVLS
GSLLRGAPYRDPLRIPREHVHNASAHLTFNKFWRRYGKLSRHLMKVKLKKYTKFLFVRDPF
VRLISAFRSKFELENEFEYRKFAVPMLRLYANHTSLPASAREAFRAGLKVSFANFIQYLLDP
HTEKLAPFNEHWRQVYRLCHPCQIDYDFVGKLETLDAAQLLQLLQVDRQLRFPPSYRNR
ASSWEEDWFAKIPLAWRQQLYKLYEADFVLFGYPKPENLLRD

Important features:

Signal peptide:

amino acids 1-31

N-glycosylation sites.

amino acids 134-137, 209-212, 280-283 and 370-373

TNFR/NGFR family cysteine-rich region protein

amino acids 329-332

FIGURE 195

TCGGGCCAGAATTCGGCACGAGGCGGCACGAGGGCGACGGCCTACGGGGCTTTGGAGGTGA
AAGAGGCCCCAGAGTAGAGAGAGAGAGACCGACGTACACGGGATGGCTACGGGAACGCGCT
ATGCCGGGAAGGTGGTGGTCTGTGACCGGGGGCGGGCGCGGCATCGGAGCTGGGATCGTGCGC
GCCTTCGTGAACAGCGGGGCCCGAGTGGTTATCTGCGACAAGGATGAGTCTGGGGGCCGGGC
CCTGGAGCAGGAGCTCCCTGGAGCTGTCTTTATCCTCTGTGATGTGACTCAGGAAGATGATG
TGAAGACCCTGGTTTTCTGAGACCATCCGCCGATTTGGCCGCCCTGGATTGTGTTGTCAACAAC
GCTGGCCACCACCCACCCACAGAGGCCTGAGGAGACCTCTGCCCAGGGATTCCGCCAGCT
GCTGGAGCTGAACCTACTGGGGACGTACACCTTGACCAAGCTCGCCCTCCCTACCTGCGGA
AGAGTCAAGGGAATGTCATCAACATCTCCAGCCTGGTGGGGGCAATCGGCCAGGCCCAGGCA
GTTCCCTATGTGGCCACCAAGGGGGCAGTAACAGCCATGACCAAAGCTTTGGCCCTGGATGA
AAGTCCATATGGTGTCCGAGTCAACTGTATCTCCCAGGAAACATCTGGACCCCGCTGTGGG
AGGAGCTGGCAGCCTTAATGCCAGACCCTAGGGCCACAATCCGAGAGGGCATGTGGCCAG
CCACTGGGCCGCATGGGCCAGCCCGCTGAGGTCGGGGCTGCGGCAGTGTTCCTGGCCTCCGA
AGCCAACTTCTGCACGGGCATTGAACTGCTCGTGACGGGGGTGCAGAGCTGGGGTACGGGT
GCAAGGCCAGTCGGAGCACCCCGTGGACGCCCCGATATCCCTTCCTGATTTCTCTCATT
CTACTTGGGGCCCCCTTCCTAGGACTCTCCACCCAAACTCCAACCTGTATCAGATGCAGC
CCCCAAGCCCTTAGACTCTAAGCCAGTTAGCAAGGTGCCGGGTCAACCCTGCAGGTTCCCAT
AAAAACGATTTGCAGCC

Figure 1 Schematic representation of the experimental design. The subjects were divided into two groups: the control group (CG) and the experimental group (EG). The CG was divided into two subgroups: the control group (CG) and the control group (CG). The EG was divided into two subgroups: the experimental group (EG) and the experimental group (EG). The CG was divided into two subgroups: the control group (CG) and the control group (CG). The EG was divided into two subgroups: the experimental group (EG) and the experimental group (EG).

<subunit 1 of 1, 270 aa, 1 stop

MATGTRYAGKVVVVVTGGGRGIGAGIVRAFVNSGARVVICDKDESGGRALEQELPGAVFILCD
 VTQEDDVKTILVSETIRRFGRLCDVNVNAGHHPPPPORPEETSAQGFRQLELNLGGTYTLTKL
 ALPYLRKSQGNVINISSLVGAIGQAQVPYVATKGAVTAMTKALALDESPYGVVNCISPGN
 IWTPLWEELAALMPDPRATIREGMLAQPLGRMGQPAEVGAAVFLASEANFCTGIELLVTGG
 AELGYGCKASRSTPVDAPDIPS

N-glycosylation site.

Short-chain alcohol dehydrogenase family protein

amino acids 10-22, 81-91, 134-171 and 176-185

FIGURE 197

AGGCGGGCAGCAGCTGCAGGCTGACCTTGCAGCTTGGCGGAATGGACTGGCCTCACAACTG
CTGTTTCTTCTTACCATTTCATCTTCCTGGGGCTGGGCCAGCCCAGGAGCCCCAAAAGCAA
GAGGAAGGGGCAAGGGCGGCCTGGGCCCCCTGGCCCCCTGGCCCTCACCAGGTGCCACTGGACC
TGGTGTCACGGATGAAACCGTATGCCCGCATGGAGGAGTATGAGAGGAACATCGAGGAGATG
GTGGCCCAGCTGAGGAACAGCTCAGAGCTGGCCCAGAGAAAGTGTGAGGTCAACTTGCAGCT
GTGGATGTCCAACAAGAGGAGCCTGTCTCCCTGGGGCTACAGCATCAACCACGACCCAGCC
GTATCCCCGTGGACCTGCCGGAGGCACGGTGCCCTGTGTCTGGGCTGTGTGAACCCCTTCACC
ATGCAGGAGGACCGCAGCATGGTGAAGCGTGCCGGTGTTTCAGCCAGGTTCTGTGCGCCGCCG
CCTCTGCCCCGCCACCGCCCCGCACAGGGCCTTGCCGCCAGCGCGAGTCATGGAGACCATCG
CTGTGGGCTGCACCTGCATCTTCTGAATCACCTGGCCCAGAAGCCAGGCCAGCAGCCCGAGA
CCATCCTCCTTGACCTTTGTGCCAAGAAAGGCCTATGAAAAGTAAACACTGACTTTTGAAA
GCAAG

3070463

FIGURE 198

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA59294

<subunit 1 of 1, 180 aa, 1 stop

<MW: 20437, pI: 9.58, NX(S/T): 1

MDWPHNLLFLLTISIFLGLGQPRSPKSKRKQGGRPGPLAPGPHQVPLDLVSRMKPYARMEEY
ERNIEEMVAQLRNSSELAQRKCEVNLQLWMSNKRSLSPWGYSINHDPRI PVDLPEARCLCL
GCVNPFTMQEDRSMVSVPVFSQVFVRRRLCPPPPRTGPCRQRAVMETIAVGCTCIF

Important features:

Signal peptide:

amino acids 1-20

N-glycosylation site.

amino acids 75-78

Homologous region to IL-17

amino acids 96-180.



FIGURE 199

GCGCCGCCAGGCGTAGGCGGGGTGGCCCTTGCGTCTCCGCTCTCCTTGAAAAACCCGGCGGG
 CGAGCGAGGCTGCGGGCCGGCCGCTGCCCCTTCCCCACACTCCCCGCGAGAGCCTCGCTCG
 GCGCCCAACATGCGGGGTGGGCGCTGCGGCCCGCAGCTAACGGCGCTCCTGGCCGCCCTGGAT
 CGCGGCTGTGGCGGCGACGCGCAGGCCCCGAGGAGGCGCGCTGCCGCCGAGCAGAGCCGGG
 TCCAGCCCATGACCGCCTCCAAC TGGACGCTGGTGATGGAGGGCGAGTGGATGCTGAAATTT
 TACGCCCCATGGTGTCCATCCTGCCAGCAGACTGATTGAGAAATGGGAGGCTTTTGCAAGAA
 TGGTGAAATACTTCAGATCAGTGTGGGGAAGGTAGATGTCATTCAAGAACAGAGTTTGTAGTG
 GCCGCTTCTTTGTCACTCTCTCCAGCATTTTTTCATGCAAAGGATGGGATATTCCGCCGT
 TATCGTGGCCCGAGGAATCTTCGAAGACCTGCAGAAATATATCTTAGAGAAGAAATGGCAATC
 AGTCGAGCCTCTGACTGGCTGGAAATCCCCAGCTTCTCTAACGATGTCTGGAATGGCTGGTC
 TTTTAGCATCTCTGGCAAGATATGGCATCTTCACAACTATTTACAGTGACTCTTGGAAAT
 CCTGCTTGGTGTTCTTATGTGTTTTTCGTATAGCCACCTTGGTTTTTGGCCTTTTTATGGG
 TCTGGTCTTGGTGGTAATATCAGAAATGTTTCTATGTGCCACTTCCAAGGCATTTATCTGAGC
 GTTCTGAGCAGAATCGGAGATCAGAGGAGGCTCATAGAGCTGAACAGTTGCAAGGATGCGGAG
 GAGGAAAAAGATGATTCAAATGAAGAAGAAAAACAAAGACAGCCTTTAGATGATGAAGAAGA
 GAAAGAAGATCTTGCGCATGAGGATGAAGCAGAGGAAGAAGAGGAGGAGGACAACTTGGCTG
 CTGGTGTGGATGAGGAGAGAAGTGAGGCCAATGATCAGGGGGCCCCCAGGAGAGGACGGTGTG
 ACCCGGGAGGAAGTAGAGCCTGAGGAGGCTGAAGAAGGCATCTCTGAGCAACCTTCCCAGC
 TGACACAGAGGTGGTGGAAAGACTCCTTGAGGCAGCGTAAAGTCAGCATGCTGACAAGGGAC
 TGTAGATTTAATGATGCGTTTTCAAGAATACACACCAAACAAATATGTCAGCTTCCCTTTGG
 CCTGCAGTTTGTACCAAATCCTTAATTTTTCTGTAATGAGCAAGCTTCTCTTAAAGATGCT
 CTCTAGTCATTTGGTCTCATGGCAGTAAGCCTCATGTATACTAAGGAGAGTCTTCCAGGTGT
 GACAAATCAGGATATAGAAAAACAAACGTAGTGTTGGGATCTGTTTGGAGACTGGGATGGGAA
 CAAGTTCATTTACTTAGGGGTGAGAGAGTCTCGACAGAGGAGGCCATTCCAGTCTTAATC
 AGCACCTTCCAGAGACAAGGCTGCAGGCCCTGTGAAATGAAGCCAAAGCAGGAGCCTTGGCT
 CCTGAGCATCCCCAAAGTGTAACGTAGAAGCCTTGATCCTTTTTCTGTGTAAAGTATTTAT
 TTTTGTCAAATTGCAAGAAACATCAGGCACCAAGTGCATGAAAAATCTTTCACAGCTAGAA
 ATTGAAAGGGCCTTGGGTATAGAGAGCAGCTCAGAAGTCATCCAGCCCTCTGAATCTCTCTG
 TGCTATGTTTTATTCTTACCTTTAATTTTTCCAGCATTTCCACCATGGGCATTGAGGCTCT
 CCACACTCTTCACTATATCTCTTGGTCAGAGGACTCCAATAACAGCCAGGTTTACATGAAC
 TGTGTTTGTTCATTCTGACCTAAGGGGTTTAGATAATCAGTAACCATAACCCCTGAAGCTGT
 GACTGCCAAACATCTCAAATGAAATGTGTGGCCATCAGAGACTCAAAAGGAAGTAAGGATT
 TTACAAGACAGATTAATAAAAAAATTGTTTTGTCCAAATATAGTTGTTGTTGATTTTTTTTT
 AAGTTTTCTAAGCAATATTTTTCAAGCCAGAAGTCTCTAAGTCTTGCCAGTACAAGGTAGT
 CTTGTGAAGAAAAGTTGAATACTGTTTTGTTTTCATCTCAAGGGGTTCCCTGGGTCTTGAAC
 TACTTTAATAATAACTAAAAAACCACTTCTGATTTTCTTTCAGTGATGTGCTTTTGGTGAAA
 GAATTAATGAATCCAGTACCTGAAAGTGAAAGATTGATTTTGTGTTCCATCTCTGTGAATC
 TTCCAAGAATTATATCTTTGTAAATCTCTCAATACTCAATCTACTGTAAGTACCCAGGAG
 GCTAATTTCTTT

00070403-101601

FIGURE 200

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56433

<subunit 1 of 1, 349 aa, 1 stop

<MW: 38952, pI: 4.34, NX(S/T): 1

MAGGRCGPQLTALLAAWIAA VAATAGPEEAA LPPEQSRVQPM TASNWTLVMEGEWMLKFYAP
WCPSCQQT DSEWEAFKNGEILQISVGKVDVIQEPGLSGRFFVTTLPAFFHAKDGIFRRYRG
PGIFEDLQNYILEKKWQSVEPLTGWKSPASLTMSGMAGLFSISGKIWHLHNYFTVTLGIPAW
CSYVFFVIATLVFGLFMGLVLVVIS ECFYVPLPRHLSERSEQNRRSEEAHRAEQ LQDAEEEEK
DDSNEEENKDSLVDDEEEKEDLGDEDAEEEEEEEDNLAAGVDEERSEANDQGPPGEDGV TRE
EVEPEEAEEGISEQPCPADTEVVEDSLRQRK SQHADKGL

Important features:

Signal peptide:

amino acids 1-22

Transmembrane domain:

amino acids 191-211

N-glycosylation site.

amino acids 46-49

Thioredoxin family proteins. (homologous region to disulfide isomerase)

amino acids 56-72

Flavodoxin proteins

amino acids 173-187

FIGURE 201

ATCTGGTTGAACTACTTAAAGCTTAATTTGTTAAACTCCGGTAAGTACCTAGCCCACATGATT
TGACTCAGAGATTCTCTTTTGTCCACAGACAGTCATCTCAGGGGCAGAAAGAAAAGAGCTCC
CAAATGCTATATCTATTACAGGGGCTCTCAAGAACAATGGAATATCATCTCTGATTTAGAAAAT
TTGGATGAAGATGGATATACTCAATTACACTTCGACTCTCAAAGCAATACAGGATAGCTGT
TGTTTTCAGAGAAAAGGATCGTGTGCTGCATCTCCTTGGCGCCTCATTGCTGTAATTTTGG
GAATCCTATGCTTGGTAATACTGGTGATAGCTGTGGTCTCTGGGTACCATGGGGGTTCTTTCC
AGCCCTTTGTCTCTTAATTGGATTATATATGAGAAGAGCTGTTATCTATTACAGCATGTCACT
AAATTCCTGGGATGGAAGTAAAAGACAATGCTGGCAACTGGGCTCTAATCTCCTAAAGATAG
ACAGCTCAAATGAATTGGGATTTATAGTAAAAACAAGTGTCTTCCCAACCTGATAATTCATTT
TGGATAGGCCTTTCTCGGCCCCAGACTGAGGTACCATGGCTCTGGGAGGATGGATCAACATT
CTCTTCTAACTTATTTAGATCAGAACCACAGCTACCCAAGAAAACCCATCTCCTAAATTGTG
TATGGATTACAGTGTGAGTCATTATGACCAACTGTGTAGTGTGCCCTCATATAGTATTTGT
GAGAAGAAGTTTTCAATGTAAAGAGGAAGGTTGAGAAGGAGAGAGAAATATGTGAGGTAGTA
AGGAGGACAGAAAACAGAAACAGAAAAGAGTAAACAGCTGAGGTCAAGATAAATGCAGAAAAATG
TTTAGAGAGCTTGGCCAACTGTAATCTTAAACCAAGAAATGAAGGGAGAGGCTGTGATTTCT
GTATTTGTGACCTACAGGTAGGCTAGTATTTATTTTCTAGTTAGTAGATCCCTAGACATGG
AATCAGGCGAGCCAAGCTTGAGTTTTTATTTTTTATTTATTTATTTTGTGAGATAGGGTCT
CACTTTGTTACCCAGGCTGGAGTGCAGTGGCACAATCTCGACTCACTGCAGCTATCTCTCGC
CTCAGCCCTCAAGTAGCTGGGACTACAGGTGCATGCCACCATGCCAGGCTAATTTTGGTG
TTTTTTGTAGAGACTGGGTTTTGCCATGTTGACCAAGCTGGTCTCTAATCTCTGGGCTTAAG
TGATCTGCCCCCTTGGCCTCCCAAAGTGCTGGGATTACAGATGTGAGCCACCACACCTGGC
CCCAAGCTTGAATTTTCATTCTGCCATTGACTTGGCATTACCTTGGGTAAGCCATAAGCGA
ATCTTAATTTCTGGCTCTATCAGAGTTGTTTTCATGCTCAACAATGCCATTGAAGTGCACGGT
GTGTTGCCACGATTTGACCTCAACTTCTAGCAGTATATCAGTTATGAACCTGAGGGTGAAAT
ATATTTCTGAATAGCTAAATGAAGAAATGGGAAAAATCTTCACCACAGTCAGAGCAATTTT
ATTATTTTCATCAGTATGATCATAATTATGATTATCATCTTAGTAAAAAGCAGGAACCTCTA
CTTTTTCTTTATCAATTAATAGCTCAGAGAGTACATCTGCCATATCTCTAATAGAACTCTT
TTTTTTTTTTTTTTTTTTTTTGTGAGACAGAGTTTCGCTCTTGTGCCCCAGGCTGGAGTGCAACGG
CAGGATCTCGGCTCACCGCAACCTCCGCCCCCTGGGTTCAAGCAATCTCCTGCCTCAGCCT
CCCAAGTAGCTGGGATTACAGTCAGGCACCACCACCCGGCTAATTTTGTATTTTTTTAGT
AGAGACAGGGTTTCTCCATGTGCGTCAGGGTAGTCCCGAACTCCTGACCTCAAGTGATCTGC
CTGCCTCGGCCCTCCCAAGTGCTGGGATTACAGGCGTGAGCCACTGCACCCAGCCTAGAATCT
TGTATAATATGTAATTGTAGGGAACTGCTCTCATAGGAAAGTTTTCTGCTTTTTAAATACA
AAAATACATAAAAAATACATAAAATCTGATGATGAATATAAAAAAGTAACCAACCTCATTGGA
ACAAGTATTAAACATTTTGGAAATATGTTTTATTAGTTTGTGTAGTACTGTTTTTCAATTTTT
ACCATTTTTTTCAGTAATTACTGTAAATGGTATTATTGGAATGAACTATATTTCTCATG
TGCTGATTGTCTTATTTTTTTCATACTTTCCCACTGGTGCTATTTTATTTCCAATGGATA
TTTCTGTATTACTAGGAGGCATTTACAGTCTCTAATGTTGATTAAATATGTGAAAGAAAT
TGTACCAATTTTACTAAATATGCACTTTAAATATGGATGATTTTATGTTATGTGGATTTTCAT
TTCAATAAAAAAAACTCTTATCAAAAAAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

00075557-101604

FIGURE 202

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA53912

<subunit 1 of 1, 201 aa, 1 stop

<MW: 22563, pI: 4.87, NX(S/T): 1

MEYHPDLENLDEDEGYTQLHFDSQSNTRIAVVSEKGS CAASEPPWRLIAVILGILCLVILVIAV
VLGTMGVLSSPCPPNWIIEKSCYLFMSLNSWDGSKRQCWQLGSNLLKIDSSNELGFIVKQ
VSSQPDNSFWIGLSRPQTEVPWLWEDGSTFSSNLFQIRTTATQENPSPNCVWIIHVSVIYDQL
CSVPSYSICEKKFSM

Important features:

Type II transmembrane domain:

amino acids 45-65

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 197-200

N-myristoylation sites.

amino acids 35-40 and 151-156

Homologous region to LDL receptor

amino acids 34-67 and 70-200.

FIGURE 205

TGCCTGGCCTGCCTTGTCACAATGCCGCTTACTCTGCTTCCAGGTTGCCCTGCCTTGCAGA
GGAAANCNTCGGGACTACACCNTCAAGTGACATGAACCTGCTGCTGGCCGTCTTCCTGCTG
GACACGAGCTTCCTGCTCAGCGNAGCCGGTGGCCCTGACAGGCTCTGAAGGCTGGCTGCCGA
GCCAGTGCCATCTTCCTGCACCTTCCTGCTCACCTGCCCTTCCTGGATGGGCCTCGAGGGG
TACAACCTCTACCGACTCGTGGTGGAGGTCTTTGGCACCTATGTCCCTGGCTACCTACTCAA
GCTGAGCGCCATGGGCTGGGGCTTCCCCATCTTTCTGGTGACGCTGGTGGCCCTGGTGGATG
TGGACAACCTATGGCCCCATCATCTTGGCTGTGCATAGGACTCCAGAGGGCGTCATCTACCCCT
TCCATGTGCTGGATCCGGGACTCCCTGGTCAGCTACATCACCACCTGGGCCTCTTCAGCCT
GGTGTTCCTGTCAACATGG

FIGURE 206

CGGACGCGTGGGCGGACGCGTGGGCGGACGCGTGGGCGGACGCGTGGGCTGGTTCCAGGTCAGGTTTTGCTTTGA
 TCCTTTTCAAAAATCGGAGACACAGAGAGGGCTCTTAGGAAAAAGTTTGGATGGGATTATGTGGAAACTACCTT
 GCGATTCTCTGCTGCCAGAGCAGGCTCGGCGCTTCCACCCAGTCAGGCTTCCCTGGCGGTGGTGAAAGAGAC
 TCGGGAGTCGCTGCTTCCAAAGTGCCCGCGTGGAGTGAGCTCTCACCCAGTCAGCCAAATGAGCCCTCTTCGGGC
 TTCTCTGCTGACATCTGCCCCGGCCGAGAGACAGGGGACTCAGGCGGAATCCAACCTGAGTAGTAAATTC
 AGTTTTCCAGCAACAAGGAACAGAACGGAGTACAAGATCCTCAGCATGAGAGAATTTACTGTGTCTACTAATG
 GAAGTATTCACAGCCACAGGTTTCTCATACTTATCCAAGAAATACGGTCTTGGTATGGAGATTAGTAGCAGTAG
 AGGAAATATGATGGATACCACTTACGTTTGATGAAAGATTGGGCTTGAAGACCAGAAGATGACATATGCAAGT
 ATGATTTTGTAGAAGTTGAGGAACCCAGTGATGGAACTATATTAGGGCCCTGGTGTGTTCTGGTACTGTACCAG
 GAAAAAGATTTCTAAAGGAAATCAAATTAGGATAAGATTGTATCTGATGAATATTTTCTTCTGAACACGGGT
 TCTGCATCCACTACAACATTTGTATGCCACAATTACAGAAGCTGTGAGTCTTTCAGTGCTACCCCTTTCAGCTT
 TGCCATCGACCTGCTTAATAATGCTATAACTGCCCTTAGTACTTTGGAAGACCTTATTCGATATCTTGAACAG
 AGAGATGGCAGTTGGACTTAGAAGATCTATATAGGCCAACTTGGCAACTTCTTGGCAAGGCTTTTGTGTTTTGAA
 GAAATCCAGAGTGGTGGATCTGAACCTTCTAACAGAGGAGGTAAAGATTATACAGCTGCACACCTCGTAACTTCT
 CAGTGTCCATAAGGGAAGAACTAAAGGAAACCGATACCATTTTCTGGCCAGGTTGTCTCCTGGTTAAACGCTGTG
 GTGGGAACGTGCTGTTGTCTCCACAATTGCAATGAATGTCAATGTGTCCCAAGCAAAGTTACTAAAAAATACC
 ACCAGGTCTTTCAGTTGAGACCAAGACCGGTGTGAGGGATTGCACAAACTACTCACCAGCTGGGCCCTGGAGC
 ACCATGAGGAGTGTGACTGTGTGTGAGAGGGAGCACAGGAGGATAGCCGATCACCACAGCAGCTCTTGCCCA
 GAGCTGTGAGTGCAGTGGCTGATTCTATTAGAGACGATATGCGTTATCTCCATCTTAACTCAGTTGTTTGCT
 TCAAGSACCTTTCATCTTCAGGATTTACAGTGCACTTCTGAAAGAGGAGACATCAACAGAAATTAGGAGTTGTGCA
 ACAGCTCTTTTGGAGAGGGCTTAAAGGACAGGAGAAAAGGTCTTCAATCTGTGGAAAGAAAATTAATGTGTGAT
 TAAATAGATCAACAGCTAGTTTTCAGAGTTACCATTGACGATTTCCACTAGCTGGGTTCTGTATTTTCAGTTCTTTC
 GATACGGCTTAGGGTAATGTGAGTACAGGAAAAAACTGTGCAAGTGAGCACCTGATTCCGTTGCTTGTCTTAAC
 TCTAAAGCTCCATGCTCTGGGCCATAAATCGTATAAAATCTGGATTTTTTTTTTTTTTTTGTCTCATATTACAT
 ATGTAACCCAGAACATTTCTATGTACTACAACTGCTGTTTTTAAAGGAACACTATGTTGCTATGAATTAACATTGT
 GTCATGCTGATAGGACAGACTGGATTTTTTCATATTTCTTATTAAATTTCTGCCATTTAGAGAGAGAACTACA
 TTTGATGTTTGGAGAGATAAACCTGAAAAGAGAGTGGCTTATCTTTCATTTATCGATAAGTCAGTTTATTGT
 TTTTATTTGTGATATTTTATATTTCTCTTTTGACATTATAACTGTGGCTTTTCTAATCTTGTAAATATATCT
 ATTTTACCAAAGGATTTTAAATTTCTTTTATGACAACTTAGATCACTATTTTAGCTTGGTAAATTTTCT
 AAACACAATTTGTTATAGCCAGAGGAACAAAGATGATATAAAATATTGTTGCTCTGACAAAAATACATGTTATTTCA
 TTCTCGTATGGTGTAGAGTTAGATTAAATCTGCATTTTAAAAAAGTGAATTTGGAATAGAATTTGGTAAAGTGCAAA
 GACTTTTGAATAAATTAATTTATCATATCTTCCATTCCTGTTATTGGAGATGAAATAAAAAGCACTTATGA
 AAGTAGACATTCAGATCCAGCCATTACCTAACCTATTCTTTTGGGGAAATCTGAGCCTAGCTCAGAAAAACAT
 AAAGCACCTTGAAGAAAGACTTGGCAGCTTCTGATAAAGCGTGTGTGCTGTGCAAGTAGGAACACATCCTATTTA
 TTGTGATGTGTGGTTTTTATTATCTTAAACTCTGTTCCATACACTTGTATATAATACATGGATATTTTTATGTACA
 GAAGTATGTCTTTAACCAGTTCACTTATTGTACTCTGGCAATTTAAAGAAAAATCAGTAAAAATTTTTGCTTGT
 AAAATGCTTTAATATNGTGCCTAGGTTATGTGGTGACTATTTGAATCAAAATGTATTGAATCATCAAAATAAAGA
 ATGTGGCTATTTTGGGAGAAAAATAAAAAAGGTTTAGGGATAACAGGTAATGCGGCC

CGGACGCGTGGGCGGACGCGTGGGCGGACGCGTGGGCGGACGCGTGGGCTGGTTCCAGGTCAGGTTTTGCTTTGA

FIGURE 207

MSLFGLLLLSALAGQRQGTQAESNLSSKFQFSSNKEQNGVQDPQHERIITVSTNGSIHSPR
FPHTYPRNTVLVWRLVAVEENVWIQLTFDERFGLEDPEDDICKYDFVEVEEPSDGTILGRWC
GSGTVPGKQISKGNQIRIRFVSDEYFPSEPGFCIHYNIVMPQFTEAVSPSVLPSSALPLDLL
NNAITAFSTLEDLIRYLEPERWQLDLEDLYRPTWQLLGKAFVFGKRSRVVDLNLLEEVRLY
SCTPRNFSVSIREEKRTDTIFWPGCLLVKRCGGNCACCLHNCNECQCVP SKVTKKYHEVLQ
LRPKTGVRGLHKSLTDVALEHHEECDVCVRGSTGG

Signal sequence:

amino acids 1-14



THE

CCGATCTCAAGCTGATCTTACGACCTTCATGCTCTGCTCTTTCACCCAGACCTTACATATCCATTTTGGAGAAG
AGCATAAAAATGGTGTTCCTTCCAAATGTGCACATGAGAGACAATAATTTATCTCTTTTAAACATATCTCAATTTCC
AAACTCTCTGGGGCTAGATGGTTTCTCTAAATACCTTCCGCTGTGATGTCACTCTGGATGTGTTCCAAAGAACCATGTG
ATCTGGGATCCGACAGACAGCAATTTGACAGAAATCTCTGGAGGTATTCCCAAGCAACACAGCAAACTCACCTCTC
ACCAATTAACCATACATACAGACATCTCCCCAGGCTCTTTTACAGACATGGACCATCTGTAGAGATCGATTTTCAGA
TGCAACTGTGTATCTATTCCACTGGGGTCAAAAACCAACATGTGCATCAAGGGCTCTGCAGATTTAAACCACCAAGG
TTTAGTGACCTCACTATTAAAACTCCTTTCTGGATGGAACACAGCTATAGAGATACCGCAGGGCTCTCCCG
CCTAGCTTACAGCTTCTCAGCTTGTAGSGCCAAACACATCTTTTCCATCAGAAAAGAGAACTTAACAGCAACTGGCT
AACATAGAAATACTCTACCTGGGCAAACTGTATTATCGAAATCTCTGTATGTGTTTCATATTCAATGAGAAA
GATGCTCTCTAAATCTGACAAAGTTTAAAGTGTCTCTCCCTGAAAGATACAATGTACACAGCGCTCCCTACTGT
TTGCCATCTTACTTTTAAACAGCAATATATCTTCAACAACAATGATTTGCAAAAATCTCAAGAGAGATGATTTTAAATAC
TCAACCAAAATTAACAAATCTTGACCTAAGTGGAAATTGCCCTCTGTGTATAATGCCCATTTCTCTGTGCGCGC
TGTAATAATAATTTCTTCCCTACAGATCCCTGTAAATGTCTTTGATGGGCTTCAGAGAAATTTAAAGTTTTCAGCTCA
CAGATAGATCTCTTTCAGCATGTGCCCCAGAGTGTTTAAAGACATCAACAACTCCAGGAATCTGGATCTGTCC
CAAAACTCTTGTGGCCAAAGAAATTTGGGGATGTCTAAATTTCTCATTTTCTCCCGAGCTCTCTCAATTTGGATCTG
TCTTTCAATTTTGAATCTCAGGCTCATCTGCGCATCTATGAACTTATCAACAGCAATTTTCTTCACTGAAAAGCTG
AAAATCTTCCGGATCAGAGGATATGTCTTTTAAAGTGTGAAAAGCTTTAACTCTTCCGCTATCATATAATCTTCAA
AATCTTGAAGTTCTTGATCTTGGCCACTAATCTTATAAAATTTGCTAACTTCAGCATGTGTTAAACAAATTTAAAGA
CTGAAAGTCATAGATCTTTTCAAGTACAATAAATTAACCTCAGAGAGATCAAGTGAAGTTGGCTCTCTGCTCAAT
CGCAGAACTCTTGTAAGAAGTTTATGAACAGCCGCTTCTGGAACATATACATTTTCAGATATGATAGATATGCA
AGGAGTTGAGCATCTCAAAAACCAAGAGGCTTCTTTCACTGTGTTAATGAAGCTGCTCAAGTATGGGCAGACCT
TTGATCTAAGTAAAAATAGTATATTTTGTCTCAAGTCTCTGATTTTCAGCATCTTTCTTTCTCTCAATGCTCT
TATCTGTGAGAAATCTATTATAGCAAATCTTAAATGGCAGTGAAATCCAACCTTTAGCAGACCTAGATATTTG
GACTTCTCCAAACAACCGGCTGTGATTTTACTCATCTCAACAGCATTTGAAGAGCTTCAACAATCGGAAGTCTCGAT
ATAAGAGCTAATAGCATTAATTTCAATCAGAGGAATATCATATATCTTAAACTTTCAGGAAGCTTAAAGGT
TCGACAGAACTGATGATGAACAGCAATGACATCTTCTCCACAGCAGGACATTCAGGAGGTGAGCTCTGTGAT
AATCTTGAATTTCAGAGGAATACCTATAGATGTTTATTGAGAGAGAGGTGATAACAGATATCTAATAATTTCAAG
AACTCTGCTAAAATTAGAGAAATAGACATCTTAAAATTCCTTAAGTTTCTTGGCTCTGGAGTTTGTATGTG
ATGCTCTCAAATCTAAAGAACTCTCTTTTGGCCAAAATGGGCTCAAACTCTTCAAGTTGGAAGAACTCTCAGTGT
CTAAAGAACTTGAAGATTTTGGACCTCAGCCACAACCTGACCACCTGCTCCAGAGATATACCAACTGCT
AGAAGCTCTCAGAATCTGATTTCTTAAAGATAATCAAATCAGGAGTCTGACGAGATTTTCTACAGATGCTCTC
CAGCTTCGATATCTGGATCTCAGCTCACTAAATAAACTCCAGATGTCAAAAAGCAGGCTCTCCAGAAATAGTCTCTC
AACAACTCGAAGATGTGCTTTTGTGATCATAAATCGGTTTCTGTGACACTGTGATGCTGTGTGGTTTGTCTGGTGG
GTTAACTACAGGAGGTGACTTATCTTCTCTGCGCCACAGATGTGACTGTGTGGGGCCAGGACACAAGAAGGC
CAAGGTGATATCTCCTGGTATCTGACACCTGTGAGTATGATCTGACTAACTGATTTCTGTTCTCACTTTCCATAT
TCGTGATCTCTCTTTCTCATGGGTGATGATGACAGCAAGTCACCTCTATTTCTGGGATGTGGGTATTTAGCAT
TTCTGTAAAGGCCAAGATAAAGGGGATCTCAGCGCTTAATATACAACAGATCTGCTATGATGCTTTATTGTGAT
GACATTAAGAGCCAGCTGTGACCGAGTGGGTTTGTGCTGAGCTGTGTGGCCAACTGTGAAGAGCCCAAGAGAA
CACTTTTAAATTTATGCTCTCAGAGAAGGGACTGGTTTACCAGGGCCAGCTGTGGAAGAACTTTCCAGAGCATAT
CAGCTTAGCAAAAAGACAGTGTGTTGTGATGACAGACAGTATGCAAGACTGAAAATTTTAAAGATGACTTTTAC
TTGTCCTCATCAGGCTCATGGATGAAAGATGTATGTGATATCTTGATATTTCTTGAAGAGCCCTTTCAGAA
TTCAGTTCTCTCGAGCTCCGGAAGAAAGCTTGTGGAGTCTGTGCTTGTAGTGGCCAAACCAACCCAGCTCAC
CCATATCTCTGCGAGTGTCTAAGAGACCGCTGGCCACAGACATATGTGGCTATAGTCAAGTGTTCAGGAA
ACAGGCTCAGCCCTCTCTTGCAAAAACCAACTGCTAGTTTACCAGAGGAGATGGCTGGC

FIGURE 209

MVFPMWTLKRQILILFNIILISKLLGARWFPKTLPCDVTLDVPKNHVIVDCTDKHLTEIPGG
IPTNTNTNLTLTINHIPDISPASFHRDLHLEIDFRCNCVPIPLGSKNMCIKRLQIKPRSFS
GLTYLKSLEYLDGNQLLEIPQGLPPSIQLLSLEANNIFSIRKENLTELANIEILYLGQNCYRR
NPCYVSYSIEKDAFLNLTCLKVLSLKDNNVTAVPTVLPSTLTELYLYNNMIAKIQEDDFNNL
NQLQILDLSGNCPCRYNAPFPCAPCKNNSPLQIPVNAFDALTELKVLRLHSNSLQHVPPRWF
KNINKLQELDLSONFLAKEIGDAKFLHFLPSLIQLDLSFNFELQVYRASMNLSQAFSSILKSL
KILRIRGYVFKELKSFNLSPLHNLQNLEVLDTGNFIKIANLSMFKQFKRLKVIDLSVNKIS
PSGDSSEVGFCSNARTSVESYEPQVLEQLHYFRYDKYARSCRFKNKEASFMSVNESCYKYGQ
TLDLKSNISFFVKSSDFQHLSFLKCLNLSGNLISQTLNGSEFQPLAELRYLDFSNNRLDLLH
STAFEELHKLEVLDISSNSHYFQSEGITHMLNFTKNLKVQLKMMNDNDISSSTSRMESES
LRTLEFRGNHLDVLWREGDNRYLQLFKNLLKLEELDISKNLSLPSGVFDGMPPNKLNLSL
AKNGLKSFSWKKLQCLKNLETLDLSHNQLTTVPERLSNCSRSILKNLILKNNQIRSLTKYFLQ
DAFQLRYLDLSSNKIQMIQKTSFPENVLNNLKMLLLHHNRFLCTCDAVWFVWVWNHTEVTIP
YLATDVTCVGPGAHKGQSVISLDLYTCELDLTNLILFSLSISVSLFLMVMMTASHLYFWDVW
YIYHFC KAKIKGYQRLISPDCCYDAFIVYDTKDPVTEWVLAELVAKLEDPREKHFNLCLEE
RDWLPQQPVLENLSQSIQLSKKTVMVTDKYAKTENFKIAFYLSHQRLMDEKVDVILIFLE
KPFQKSKFLQLRKRLCGSSVLEWPTNPQAHFYPFWQCLKNALATDNHVAYSQVFKETV

Signal sequence:

amino acids 1-26

Transmembrane domain:

amino acids 840-860

FIGURE 210

GGGTACCATTCTGCGCTGCTGCAAGTTACGGAATGAAAAATTAGAACAACAGAAACATGGAAAAATGTTCTCTTC
 AGTCGTCAATGCTGACCTGCATTTTCTGCTAATATCTGGTTCCTGTGAGTTATGCGCCGAAGAAAAATTTTCTA
 GAAGCTATCCTTGTGATGAGAAAAAGCAAAATGACTCAGTTATTGCGAGAGTGCAGCAATCGTGCATACAGGAAG
 TTCCCAACACGGTGGGCAAAATAGTGACAGAACTAGACCTGTCTGATAAATTCATCAACACATAACGAATGAAT
 CATTTCAAGGGCTGCAAAATCTCATTAAATAAAATCTAAACCACAAACCCAAATGTACAGCACCAGAGACGGAAATC
 CCGGTATACAATCAAATGGGCTTGAATTCACAGACGGGGCATTCTCAACTAAAAACCTAAGGGAGTTTACTGCT
 TTGAAGACACACAGTTACCCCAAAATACCTCTGCTGGTTTGCAGAGTCTTTTGACAGAACTTAGTCTAATTCAAACA
 ATATATACAACATAAATTAAGAGGGGCAATTTCAAGACTTATAAACTTGAAAAATCTCTATTGGGCTGGAACTGCT
 ATTTTAAACAAGATTTGCGAGAAAACTAACATAGAAGATGGAGTATTGGAAACCGTGCACAAATTTGGAGTTGCTAT
 CACTATCTTTCAATTCTCTTTACACGTCGCCACCAAACTGCCAAGCTCCCTACGCAAACTTTTTCTGAGCAACA
 CCCGAGTCAAATACATTAGTGAAGAAGATTTCAGGGGATTGATAAATTTAAACATTACTAGATTAAAGCGGAACT
 GTCCGAGGTGCTCTCAATGCCCAATTTTCATGCGTGCTTTGATGGTGGTGTCTCAATTAATATAGATCGTTTGT
 CTTTTCAAAACTTGACCCAACCTCGATCACTAAACCTCTCTAGCACTTCCTCTCAGGAAGATTAACTGCTGCTGGT
 TTAAAAATATGCTCATCTGAAGTGTCTGGATCTTGAATTCAACTATTATTAGGGGAGAAATAGTCTCTGGGGCAT
 TTTTAAACGATGCTGCCCGCTTAGAATACTTGACTGTCTTTTAACTATATAAAGGGGAGTTATCCACAGCAAT
 TTAATATTTCCAGAAATCTCTCTAAACTTTTGCTCTACGGGCATTGCAATTAAGAGGTTATGCTTCCAGGAAC
 TCAGAGAAATGATGATTTCGACCCCTTAGATGCAGTCTCCAACTTATGCAATCACTAAGCTTGGGTATTAATTTTATTA
 AGCAAACTCGAATTCAAACTTTTCAAATTTTCCAACTTCCAACTCGGAATAATTACTTGTGCAGAAAAACAGAAATCAC
 CGTTGGTAAAGATACCCGGCAGAGTTATGCAAAATAGTTCCTCTTTTCAACGTCATATCCGGAACACGCGCTCAA
 CAGATTTTGAGTTTGACCCCACTTCGAACCTTTTATCATTTCAACCGCTCTTTTATAAAGCCCAATGTGCTGCTT
 ATGGAAGAAGCCTTAGATTAAAGCCTCAACAGTATTTTCTTATTGGGCCAAACCAATTTGAAATCTTCTCGAACA
 TTGCTCTGTTTAAATCTGCTCTGCAAAATAGCAATGCTCAAGTGTAAAGTGGAACTGAAATTTTCAGCCATTCCTCATG
 TCAAAATATTGGATTTCGACAAACAATAGACTAGACTTTGATAATGCTAGTGCTCTTAGTGAATTTGCCAGTTGG
 AAGTTCTAGATCTCAGCTATAATTCACATATTTCAAGATAGCAGGCGTAACACATCATCTAGAAATTTATTCAAA
 ATTTCAAAAATCTAAAAGTTTAAACTTGAGCCACACACAACTTATACTTTAAACAGATAAGTATTAACCTGGAAA
 GCAAGTCTCGGTGAAATTTAGTTTCAGTGGCAATCGCCTTGACATTTTGTGGAATGATGATGACAAACAGGTATA
 TCTCCATTTTCAAAGGCTCAAGAATCTGACACGCTTGGATTATTCCTTAAATAGGCTGAAGCACATCCCAATG
 AAGCATTCCTTAATTTGGCCAGCGAGTCTCACTGAACATACATATAAAGTAAATATGTTAAAGTTTAACTGGA
 CATCTCCAGCAAGCTTCTCGCTCTCAGTTGCTTGACTTACGTGGAACCAAACTACTCTTTTAACTAGTAGCC
 TATCTGACTTTTACATCTTCCCTCTGGACACTGCTGCTGAGTCTAAGCAGGATTTCCCACTACCTCTGCTCTT
 TTTCTGAAGTCAGTGTCTGAAGCACCTCGATTAAAGTTCCAATCTGCTAAACCAATCAACAAATCCGCACTTG
 AAACTAAGACACCAACCAATTTATCTATGTTGGAATACACGGAACCCCTTTGAATGCACCTGTGACATTTGGAG
 ATTTCCGAAGATGGATGGATGAACATCTGAATGTCAAATTCAGACTGGTATGATGATTTTGTGCCAGTCTCTG
 GGGATCAAAGAGGGAAGAGTATTGTGAGTCTGGAGCTAAACACTTGTGTTTTCAGATGTCACTGCAGTGATATAT
 TTTTCTTCAAGTCTTTTATCACCACTGGTTATGTTGGCTGCCCCTGCTCAACATTTGTTTAACTGGGATGTTT
 GGTTTATATATTAATGTGTGTTTAGCTAAGGTAAGGCTACAGGTCCTTTTCCACATCCCAAACTTTCTATGATG
 CTTCATCTTTCTTATGACACCAAGAGTCTCTGTTACTGACTGGGTGATAAATGAGCTGGCTTACCACCTTGAAG
 AGAGCCGAGACAAAACGTTCTCTTTGTCTAGAGGAGAGGAGTGGGACCCGGGATTTGGCCATCATCGACAAC
 TCATGACAGGATCAACCAAGCAAGAAAAAGTATTTGTTTAAACCAAAAAATATGCAAAAAGCTGGAATTTA
 AAACAGCTTTTACTTGGCTTTGAGAGGCTAATGGATGAGAACATGGATGTGATTATATTTATCTCTGCTGGAGC
 CAGTGTACAGCATTCTCAGTATTTGAGGCTACGGCAGCGGATCTGTAAGAGCTCCATCTCTCAGTGGCCTGACA
 ACCCAAGGCAAGGCTTGTTTTGGCAAACTCTGAGAAATGTGCTCTTGAAGTCAAAATGATTACAGGTATAACA
 ATATGATGTCGATTTCATTAAAGCAATCTAACTGACGTTAAGTCATGATTTCGCGCAATTAAGAGTGAAG
 GAATGACATTTCTGTATTAGTTATCTATTGCTATGTAAACAAATATCCCAAACTTAGTGGTTTAAACAACACA
 TTTGCTGGCCACAGTTTTTGAGGCTCAGGAGTCCAGGCCAGCATACCTGGTCTCTGCTCAGGGTGTCTCAG
 AGGCTGCAATGTAGGTGTTTCAACAGAGCATATGGCATCTAGTGGGTGACCATGATGTGGTTGTTTCTGGATTAT
 ATTCTCTCTGGGCTATTGGCCAAAGGCTATACTCATGTAAGCCATGCGAGCTCTCCCAACAGGACGCTGTCTT
 ATCAGAGCTAGCAAAAGAGAGGTTGCTAGCAAGTGAAGTCAAACTTTTGAATCGAATCAAAAAGTGAAT
 ATCTCATCTTTTGGCCATTTCTATTGTTGTAAGTAACCAACAGGTCACAGGATCCAGGAGTGAACCAT
 TCAGTCCAGGGAAAAACGCTGAAGACCAAGATGGTGAAGTCTGATTGCTTCAGTTGCTCATCAACTATTTTCCCT
 TGACTGTGCTGCTGGGATGGCCTGCTATCTTGATGATAGATTGTGAATATCAGAGGCGAGGGATCACTGTGGACC
 ATCTTAGCAGTTGACCTAACACATCTTCTTTTCAATATCTAAGAACTTTTGGCAATGTGACTAATGGTCTTAATA
 TTAAGTGTGTTGTTATATTTATCATATCTATGCTATGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT
 TTAGCTGTCTTTTCAAAATATTGCTGTAAACATTTGACTTCAAGGTTTAGATGCAATTTAAGAACTGATGAG
 ATAGCTTTTAAAGCATCTTTTACTTCTTACCATTTTAAAAAGTATGACGCTAAATTCGAGGCTTTTGGTCTATA
 TTGTTAATTTGCCATTGCTGTAATCTTAAATGAATGAATAAATGTTTCATTTTCAAAAAA

102101 102101 102101

FIGURE 212

CCAGGTCCAAC TGACCTCGGTTCTATCGATTGAAT TCCCCGGGGATCCTCTAGAGATCCCT
CGACCTCGACCCACGCTCCGCCAAGCTGGCCCTGCACGGCTGCAAGGGAGGCTCCTGTGGA
CAGGCCAGGCAGGTGGGCCTCAGGAGGTGCCTCCAGGCGGCCAGTGGGCCTGAGGCCCCAGC
AAGGGCTAGGGTCCATCTCCAGTCCCAGGACACAGCAGCGGCCACCATTGGCCACGCCTGGGC
TCCAGCAGCATCAGCAGCCCCCAGGACCGGGAGGCACAGGTGGCCCCCACCACCCGGAGGA
GCAGCTCCTGCCCTGTCCGGGGGATGACTGATTCTCTCTCCGCCAGGCCACCAGAGGAGAA
GGCCACCCCGCTTGAGGCACAGGCCAATGAGGGGCTCTCAGGAGGTGCTGCTGATGTGGCTT
CTGGTGTGGCAGTGGGCGGCACAGAGCACGCCTACCGGCCCGCCGTAGGGTGTGTGCTGT
CCGGGCTCACGGGGACCCGTCTCTCCGAGTCGTTCTGTGCAGCGTGTGTACCAGCCCTTCCTCA
CCACCTGCGACGGGCACCGGGCTGCAGCACCTACCGAACCATCTATAGGACCGCCTACCGC
CGCAGCCCTGGGCTGGCCCTGCCAGGCCTCGCTACGCGTGTGCCCCGGCTGGAAGAGGAC
CAGCGGGCTTCTGGGGCTGTGGAGCAGCAATATGCCAGCCGCATGCCGGAACGGAGGGA
GCTGTGTCCAGCCTGGCCGCTGCCGCTGCCCTGCAGGATGGCGGGGTGACACTTGCCAGTCA
GATGTGGATGAATGCAGTGTCTAGGAGGGGCGCTGTCCCAGCGCTGCATCAACACCGCCGG
CAGTTACTGTGTGCCAGTGTGTGGGAGGGGCACAGCCTGTCTGCAGACGGTACACTCTGTGTGC
CCAAGGGAGGGCCCCCAGGGTGGCCCCCAACCCGACAGGAGTGGACAGTGCATGAAGGAA
GAAGTGCAGAGGCTGCAGTCCAGGGTGGACCTGCTGGAGGAGAAGCTGCAGCTGGTGTGGC
CCCACTGCACAGCCTGGCCTCGCAGGCACTGGAGCATGGGCTCCCGACCCGGCAGCCTCC
TGGTGCATCCTTCCAGCAGCTCGGCCGCATCGACTCCCTGAGCGAGCAGATTTCTTCTCTG
GAGGAGCAGCTGGGGTCTGCTCCTGCAAGAAAGACTCGTGACTGCCCAGCGCCCCAGGCTG
GACTGAGCCCTCACGCCGCTGCAGCCCCATGCCCTGCCCAACATGCTGGGGTCCAG
AAGCCACCTCGGGGTGACTGAGCGGAAGGCCAGGCAGGGCCTTCTCTCTTCTCTCTCCCC
TTCTCTCGGGAGGCTCCCCAGACCCTGGCATGGGATGGGCTGGGATCTTCTCTGTGAATCCAC
CCCTGGCTACCCCCACCTGGCTACCCCAACGGCATCCCAAGGCCAGGTGGGCCCTCAGCTG
AGGGAAGGTACGAGCTCCCTGTCTGGAGCCTGGGACCCATGGCACAGGCCAGGCAGCCCCGGAG
GCTGGGTGGGGCCTCAGTGGGGGCTGTGCCTGACCCCCAGCACAAATAAAATGAAACGTGA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGGGCGGCCGCACTCTAGAGT
CGACCTGCAGAAGCTTGGCCGCATGGCCCAACTTGTTTATTGCAGCTTATAATGGTTACAAAT

FIGURE 213

MRGSQEVLLMWLLVLAVGGTEHAYRPGRRVCAVRAHGDPVSESFVQRVYQPFLTTCDGHRAC
STYRTIYRTAYRRSPGLAPARPRYACCPGWKRTSGLPGACGAAICQPPCRNGGSCVQPGRCR
CPAGWRGDTQCSDVDECSARRGGCPQRCINTAGSYWCQCWEGHSLSADGTLCVPKGGPPRVA
PNPTGVDSAMKEEVQRLQSRVDLLEEKQLQLVLAPLHSLASQALEHGLPDPGSLLVHSFQQLG
RIDSLSEQISFLEEQLGSCSCKKDS

Signal sequence:

1-19

FIGURE 214

GCCAGGCAGGTGGGCCTCAGGAGGTGCCTCCAGGCGGCCAGTGGGCCTGAGGCCCCAGCAAG
GGCTAGGGTCCATCTCCAGTCCCAGGACACAGCAGCGGCCACCATGGCCACGCCTGGGCTCC
AGCAGCATCAGAGCAGCCCCCTGTGGTTGGCAGCAAAGTTAGCTTGCTGGGCCCCGCTGTGA
GGGGCTTCGCGCTACGCCCTGCGGTGTCCCGAGGGCTGAGGTCTCCTCATCTTCTCCCTAGC
AGTGGATGAGCAACCCAAACGGGGGCCCGGGGAGGGGAAC TGCCCCGAGGGAGAGGAACCC
AAAGCCACATCTGTAGCCAGGATGAGCAGTGTGAATCCAGGCAGCCCCCAGGACCGGGGAGG
CACAGGTGGCCCCCACCACCCGGAGGAGCAGCTCCTGCCCTGTCCGGGGATGACTGATTCT
TCTCCGCCAGGCCACCCAGAGGAGAAGGCCACCCCGCTGGAGGCACAGGCCATGAGGGGC
TCTCAGGAGGTGCTGCTGATGTGGCTTCTGGTGTGGCAGTGGGCGGCACAGACACGCCTA
CCGGCCCGCGCTAGGGTGTGTGCTGTCCGGGCTCACGGGGACCTGTCTCCGAGTCGTTCTG
TGCAGCGTGTGTACCAGCCCTTCTCACCACCTGCGACGGGCACCGGCCTGCAGCACCTAC
CGAACCATCTATAGGACCGCCTACCGCCGAGCCCTGGGCTGGCCCTGCCAGGCCTCGCTA
CGCGTGTGCCCCGGCTGGAAGAGGACCAGCGGGCTTCTGGGGCTGTGGAGCAGCAATAT
GCCAGCCGCCATGCCGAACGGAGGGAGCTGTGTCCAGCCTGGCCGCTGCCGCTGCCCTGCA
GGATGGCGGGGTGACACTTGCCAGTCAGATGTGGATGAATGCAGTGCTAGGAGGGGCGGCTG
TCCCCAGCGTGCATCAACACCGCCGGCAGTTACTGTTGCCAGTGTGTGGAGGGGCACAGCC
TGCTGTCAGACGGTACACTCTGTGTGCCCAAGGAGGGCCCCCAGGGTGGCCCCCAACCCG
ACAGGAGTGGACAGTGCAATGAAGGAAGAAGTGACAGAGCTGCAGTCCAGGGTGGACCTGCT
GGAGGAGAAGCTGCAGCTGGTGCTGGCCCCACTGCACAGCCTGGCCTCGCAGGCACTGGAGC
ATGGGCTCCCGGACCCCGGCAGCCTCCTGGTGCACTCCTTCCAGCAGCTCGGCCGCATCGAC
TCCCTGAGCGAGCAGATTTCTTCTTGAGGAGCAGCTGGGGTCTGCTCCTGCAAGAAAGA
CTCGTGA CTGCCCAGCGCTCCAGGCTGGACTGAGCCCCCTACGCCGCCCTGCAGCCCCCATG
CCCCTGCCCAACATGCTGGGGGTCCAGAAGCCACCTCGGGGTGACTGAGCGGAAGGCCAGGC
AGGGCCTTCCTCCTCTTCTCCTCCTCCCCCTTCTCGGGAGGCTCCCCAGACCTGGCATGGGAT
GGGTGGGATCTTCTCTGTGAATCCACCCCTGGCTACCCCCACCCTGGCTACCCCAACGGCA
TCCCAAGGCCAGGTGAGCCCTCAGCTGAGGGAAGGTACGAGCTCCCTGCTGGAGCCTGGGAC
CCATGGCACAGGCCAGGCAGCCCGAGGCTGGGTGGGGCCTCAGTGGGGCTGCTGCCTGAC
CCCCAGCACATAAAAAATGAAACGTG

[illegible]

Signal sequence:

1-19

FIGURE 216

CCCACGCGTCCGAAGCTGGCCCTGCACGGCTGCAAGGGAGGCTCCTGTGGACAGGCCAGGCA
GGTGGGCCCTCAGGAGGTGCCTCCAGGCGGCCAGTGGGCCCTGAGGCCCCAGCAAGGGCTAGGG
TCCATCTCCAGTCCCAGGACACAGCAGCGGCCACCATGGCCACGCCTGGGCTCCAGCAGCAT
CAGCAGCCCCCAGGACCGGGGAGGCACAGGTGGCCCCCACCACCCGAGGAGCAGCTCCTGC
CCCTGTCCGGGGGATGACTGATTCTCCTCCGCCAGGCCACCCAGAGGAGAAGGCCACCCCGC
CTGGAGGCACAGGCCATGAGGGGCTCTCAGGAGGTGCTGCTGATGTGGCTTCTGGTGTGGC
AGTGGGCGGCACAGAGCACGCCTACCGGCCCGGCCGTAGGGTGTGTGCTGTCCGGGCTCACG
GGGACCTGTCTCCGAGTCGTTCTGTGCAGCGTGTGTACCAGCCCTTCTCACCACCTGCGAC
GGGCACCGGCCCTGCAGCACCTACCGAACCATCTATAGGACCGCCTACCGCCGACGCCCTGG
GCTGGCCCTGCCAGGCCTCGCTACGCGTGTGTGCCCCGGCTGGAAGAGGACCAGCGGGCTTC
CTGGGGCCTGTGGAGCAGCAATATGCCAGCCGCCATGCCGAACGGAGGGAGCTGTGTCCAG
CCTGGCCGCTGCCGCTGCCCTGCAGGATGGCGGGGTGACACTTGCCAGTCAGATGTGGATGA
ATGCAGTGCTAGGAGGGGCGGCTGTCCCCAGCGCTGCGTCAAACCCGCGGCAGTTACTGGT
GCCAGTGTGGGAGGGGCACAGCCTGTCTGCAGACGGTACACTCTGTGTGCCAAGGGAGGG
CCCCCAGGGTGGCCCCCAACCCGACAGGAGTGGACAGTGCAATGAAGGAAGAAGTGCAGAG
GCTGCAGTCCAGGGTGGACCTGCTGGAGGAGAAGCTGCAGCTGGTGTGTGCCCCACTGCACA
GCCTGGCCTCGCAGGCACCTGGAGCATGGGCTCCCGGACCCGGCAGCCTCCTGGTGCCTCC
TTCCAGCAGCTCGGCCGCATCGACTCCCTGAGCGAGCAGATTTCTTCTCTGGAGGAGCAGCT
GGGGTCTCTCTCTGCAAGAAAGACTCGTGACTGCCCAGCGCCCCAGGCTGGACTGAGCCCC
TCACGCCGCCCTGCAGCCCCCATGCCCCCTGCCAACATGCTGGGGGTCCAGAAGCCACCTCG
GGTGACTGAGCGGAAGGCCAGGCAGGGCCTTCCTCCTCTTCTCCTCCTCCCTTCTCCTGGGAG
GCTCCCCAGACCTTGGCATGGGATGGGCTGGGATCTTCTCTGTGAATCCACCCTTGGCTACC
CCCACCTTGGCTACCCCAACGGCATCCCAGGCCAGGTGGGCCCTCAGCTGAGGGAAGGTAC
GAGCTCCCTGCTGGAGCCTGGGACCATGGCACAGGCCAGGCAGCCCGAGGCTGGGTGGGG
CCTCAGTGGGGGTGCTGCCTGACCCCCAGCACAAATAAAATGAAACGTG

FIGURE 217

MRGSQEVLLMWLLVLAVGGTEHAYRPGRRVCAVRAHGDPVSESFVQRVYQPFLTTCDGHRAC
STYRTIYRTAYRRSPGLAPARPRYACCPGWKRTSGLPGACGAATCQPPCRNGGSCVQPGRCR
CPAGWRGDTQCSDVDECSARRGGCPQRCVNTAGSYWCQCWEGHSLSADGTLCVPKGGPPRVA
PNPTGVDSAMKEEVQRLQSRVDLLEEKQLQLVLAPLHSLASQALEHGLPDGSLLVHSFQQLG
RIDSLSSEQISFLEEQLGSCSCKKDS

Signal sequence:

1-19

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

GGTTGCCACAGCTGGTTTAGGGCCCGCCACAGCTGGGGCCGCTTGTTCAGGAGGAGACAGCTCCCGGCCCGGGGAG
GACAGATCGCTGCCACCTTTGGCTCGCCGACGTGATTCCTCGGACGGTCGTTTCTCGCGTCAGCTGCGCGCGCG
AGTTGGGTCTCCGTGTTTCAGGCCGGGCTCCCCCTTCTGGTCTCCCTTCTCCGCTGGGCGCGTTTATCGGGAGG
AGATTGTTCTTCAGGGCTAGCAATTGCACTTTTGATGATGTTTGACCCAGCGGAGAAATAGCAGCAACCGTGAT
TTCAAAGCTGGGCTCAGCCTCTGTTTCTTCTCTGTAAATCGCAAAACCAATTTTGGAGCAGGAATTCAAATCA
TGTCTGTGATGTGTGAGAAGAAGGTGACAGCAAAATGGGAGAAATCCCAAGCAGGAACCACTTGTCTGTG
ATGGCCGCGTCATGTAGCCCGCGGCAAGAGGCACTTTTCTACTGACCCCTTTCTCATCTCTGGGACATGTGACAT
TCTTCTTCGCCCTTTGAGTGC CGCTACCTGGCTGTTCAGCTGTCTCTCGCATCCCTGTATTTTGTGCCATGCTCT
TCCCTTTCTCATGGTCTACCTTGTGAGGACCACTTCAGTACCCCTGGAGTATTTCTGGGCGCTACACAGAT
AAGCACTTTTCATAGAAATGAGATGATGAAGTACCAATAGTGTGGCTGCCCGAGGCGCAGCAACGCCCTGTGA
TTCAGAATTTTCAGATAAACAACAGATTTGAAATCAATAACTGTATACACATCAAGATCTCTCCGCTCTCCCC
GGGCTCCCAATTGTCAGCATCTGTGACAACTGTGTGGAGCGCTTGACCACTACTGCCCTGGGTGGGGAATTTGT
TTGAAAAGAGAACTACCGCTACTTTACCTCTTCACTCTTCTCTCCCTCTCAATATATGTCTTCGCCCT
TCAACATCTGTCTATGTGGCCCTCAAATCTTTGAAATTTGGCTTTCTGGAGACATGGAAGAAACTCTGGAACATG
TTTCAGAAGTCTCATTTTGTCTTTTACACTCTGTGTCGTGCGTGCGGATGACATGGAATTTCACTACTTCTCTGTG
CTCTCAACAGACAGCAACCAATGAGCATGACAAAGGATCATGACAGGGAAGAATCGCGTCAAGCTCCCTACAGCC
ATGGCAATATTTGGAAGAATCTGTGAAGTCTGTGTGGAGCTTGGCCCCGATGTGTGATCGAAGGGGTAT
TTTTGCCACATGGGAGAAATGGAAATGCACCTCCAGATCAAGAGACAGTACAGCCTTGTGCCACAGCGC
AGCCCTCCAGCAAACTCAATCAATGAGATGCGCGAGGACAGCAGCATCCCGAAGAGATGCCCACTTCAG
AGCCCCCAGAGCCACCAAGAGGCGAGCTGAAGCTGAGAAGTAGCCCTATCTATGGAAGAGACTTTTGTTTGTGTT
TAATTAGGGCTATGAGAGATTTCAGGTGAGAAGTTAAACTCTGAGACAGAGACAGTAAAGCTGTCCCTTTTAACT
GTATTTCTTTGGTCTTTTAACTCACCAGTGTGCACATGCACTTTCTCTGTCGAAGCTTTTAAATTTCTGAAC
CAAGGCACTGGCAGAGATGTCAGTCACTCTGTATGATCAAGAAAATGGGTCTCTTGGGCCCTGGCATGGTCTCT
CCATGGCTCAGCAGCGGTTCCCTTGAGACCCCTCTCTTCCCTCAGATCCGAGCCCTCTCTGTTGGGTCA
TGCTCTCATTTCTGGGGCTAAAGATTTTTGAGATGGCTCAAACTCTCCCAAGCTCTGCACGTGCTGAGTCAAG
GGCAGTCAAGAGACCTCTGGCCAGGGGATCTCAACTGGTCTTGGGGTCTTCAGGACTGAAGAGAGGGAG
TGGGGTCAGAAGATTCTCTGGCCACAGTGCACAGATGCCCCAAATCTTTAGGAATGGGACAGGTACCT
TCACATTTGTATTANNNNNNNNNNNNNNNNNNNNNNNNNNNNNTTTTCTTTCTTGTACTCTGCTCCATATGAG
CAGGAATGGCGATGATAAAAGATCTGCATTTGGTCAATTTCTTTCTTCAGAGAAGCCGAGTGTGACTATTAAG
ACTATCCCTCAGACTCCCTGTGTGAGGCTCGAGAGGCTCGTAAGTGCACAAATGGGAAACCAAGGCACAGAG
GCTCTCCTCTCTCTCTCTCTCCCGGATGACCTCAAAAAAAAAATGCTTACCAAGCTTCTCTATTAAGCT
CGSCTGAGTGAGGGAAGCCAGCACTGCTGCCCTCTCGGGTAACTCACTTCAAGGCTCGGCCCACTCTGGCT
ATGGTAAACCACTCGGGGCTTCTCAGGCCCGCTCTCCAGCACTCCACCGGAGATCCGAGGCCATCT
CACCTTGGGGGTGGGCTGTGGCCCGCAGTCACTCTGTCAGGACCTGTCTTAATTTCAGGGAAGAGGATTTATGT
ATTATATGTGGCTATATTTCTCAGACAGCACTGTGTTTCTCTTTCTCAAGCCAGGGTCTCTGTGGATGACTTAT
CGGGTGGGGGAGTGAAACCGGAATTTTCACTATTTTGAAGCGGATTAACGTGTCTAATGCA

FIGURE 219

MSVMVVRKKVTRKWEKLPGRNTFCCDGRVMMARQKGIFYLTFLILGTCTLFFAFECRYLAV
QLSPAIPVFAAMLFLFSMATLLRTSFSDPGVIPRALPDEAAFIEMEIEATNGAVPQGQRPPP
RIKNFQINNQIVKLKYCYTCKIFRPPRASHCSICDNCVERFDHHCWPVGNVCVKRNYRYFYFYL
FILSLSLTIYVFAFNIVYVALKSLKIGFLETLPKETPGTVLEVLICFFTLWSVVGLTGFHTF
LVALNQTTNEDIKGSWTGKNRVQNPYSHGNIVKNCCEVLCGPLPPSVLDRRGILPLEESGSR
PPSTQETSSSLLPQSPAPTEHLNSNEMPEDSSTPEEMPPPEPPEPPQEAAEAEK

Putative transmembrane domains:

amino acids 36-55 (type II TM), 65-84, 188-208, 229-245

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

AAAACCCCTGTATTTTTTACAATGCAATAGACAATNANCTGGAGGCTTTTGAATTAGGTAT
 TATAGGGATGGTGGGGTTGATTTTTNTTCTCGAGGCTTTTGGCTTTGGACTCTCNCITTTCT
 CCCACAGAGCNCITTCGACCATCACTGCCCTGGGTGGGAATTGTGTTGGAAAGAGGAAC
 CCGCTANTTCTACCTCTTCATCCTTTNTCTCTCCNCCTACAATCTATGTCTTCGCCTTCA
 ACATCGT

FIGURE 221

GTGTGTCTTCAGCAAAACAGTGGATTAAATCTCCTTGCACAAGCTTGAGAGCAACACAA
TCTATCAGGAAAGAAAGAAAGAAAAAACCGAACCTGACAAAAAGAAGAAAAAGAAGA
AAAAAATCATGAAAACCATCCAGCCAAAAATGCACAATTCTATCTCTTGGGCAATCTTCAC
GGGGCTGGCTGCTCTGTGTCTCTTCCAAGGAGTGCCCGTGCGCAGCGGAGATGCCACCTTCC
CCAAAGCTATGGACAACGTGACGGTCCGGCAGGGGAGAGCGCCACCCTCAGGTGCACCTATT
GACAACCGGGTCACCCGGGTGGCCTGGCTAAACCGCAGCACCATCCTCTATGCTGGGAATGA
CAAGTGGTGCCTGGATCCTCGCGTGGTCTTCTGAGCAACCCAAACGCAGTACAGCATCG
AGATCCAGAACGTGGATGTGTATGACGAGGGCCCTTACACCTGCTCGGTGCAGACAGACAAC
CACCCAAAGACCTCTAGGGTCCACCTCATTGTGCAAGTATCTCCAAAATTGTAGAGATTTCT
TTCAGATATCTCCATTAAATGAAGGGAACAATATTAGCCTCACCTGCATAGCAACTGGTAGAC
CAGAGCCTACGGTTACTTGGAGACACATCTCTCCAAAGCGTTTGGCTTTGTGAGTGAAGAC
GAATACTTGGAAATTAGGGCATCACCCGGGAGCAGTCAGGGGACTACGAGTGCAGTGCCTC
CAATGACGTGGCCGCGCCCGTGGTACGGAGAGTAAAGGTACCCGTGAACCTATCCACCATACA
TTTCAGAAGCCAAGGGTACAGGTGTCCCGTGGGACAAAAGGGGACACTGCAGTGTGAAGCC
TCAGCAGTCCCTCAGCAGAATTCCAGTGGTACAAGGATGACAAAAGACTGATTGAAGGAA
GAAAGGGGTGAAAGTGGAAAACAGACCTTTCTCTCAAAACTCATCTTCTTCAATGTCTCTG
AACATGACTATGGGAACTACACTTGCCTGGCCTCCAACAAGCTGGGCCACACCAATGCCAGC
ATCATGCTATTGGTCCAGGCGCCGTCAGCGAGGTGAGCAACGGCACGTGAGGAGGGCAGG
CTGCGTCTGGCTGCTGCCTCTTCTGGTCTTGACCTGCTTCTCAAATTTTGAATGTGAGTGCC
ACTTCCCCACCCGGGAAAGGCTGCCGCCACCACCACCAACAACAAGCAATGGCAACAC
CGACAGCAACCAATCAGATATATACAAATGAAATTAGAAGAAAACACAGCCTCATGGGACAGA
AATTTGAGGGAGGGGAACAAAGAATACTTTGGGGGGAAAAGAGTTTTAAAAAGAAATTGAA
AATTGCCTTGAGATATTTAGGTACAATGGAGTTTTCTTTCCCAAACGGGAAGAACACAGC
ACACCCGGCTTGACCCACTGCAAGCTGCATCGTGCAACCTCTTTGGTGCCAGTGTGGGCAA
GGGCTCAGCCTCTCTGCCACAGAGTGCCCCACGTGGAACATTCTGGAGCTGGCCATCCCA
AATTCAATCAGTCCATAGAGACGAACAGAATGAGACCTTCCGGCCCAAGCGTGGCGCTGCGG
GCACTTTGGTAGACTGTGCCACCACGGCGTGTGTGTGAAACGTGAAATAAAAAGAGCAAAA
AAAAA

THE

Signal peptide:

amino acids 1-28

[illegible]

GAAAAAAATCATGAAAACCATCCAGCCAAAAATGCACAATTCTATCTCTTGGGCAATCTTC
ACGGGGCTGGCTGCTCTGTGTCTCTTCCAAGGAGTGCCCGTGCAGCGGAGATGCCACCTT
CCCCAAAGCTATGGACAACGTGACGGTCCGGCAGGGGGAGAGCGCCACCTCAGGTGCACTA
TTGACAACCGGGTCAACCGGTGGCTGGCTAAACCGCAGCACCATCTCTATGCTGGGAAT
GACAAGTGGTGCCTGGATCTCGCTGGTCTCTTCTGAGCAACACCCAAACGCAGTACAGCAT
CGAGATCCAGAACGTGGATGTGTATGACAGGGGCCCTTACACCTGCTCGGTGCAGACAGACA
ACCACCCAAAGACCTCTAGGGTCCACCTCATTGTGCAAGTATCTCCCAAATTTAGAGATT
TCTTCAGATATCTCCATTAATGAAGGGAACAATATTAGCCTCACCTGCATAGCAACTGGTAG
ACCAGAG

FIGURE 224

ATGGCTGGTGACGGCGGGGCCGGGCAAGGGACCGGGCCGCGGCCGGGAGCGGGCCAGCTGCGGGAGCCCTGA
ATCACCGCTTGGCCCCGACTCCACCA~~ATGA~~ACGTCGCGCTGCAGGAGCTGGGAGCTGGCAGCAACGTTGGGATTCAG
AAGGGGACAAGACAGCTGTTAGGCTCACGCACGCAGCTGGAGCTGGTCTTAGCAGGTGCCTCTCTACTGCTGGCT
GCACTGCTTCTGGGCTGCTTGTGGCCCTAGGGGTCCAGTACCAAGAGAGCCCATCCACAGCACCTGCTTACA
GAGGCTGCACTCGAGTGGCTGGAAAAATCCTGGATCCCTGGACCGAGGGGTGAGCCCCCTGTGAGGACTTTTAC
CAGTTCTCTGTGGGGCTGGATTCTGGAGGAACCCCTGCCGATGGGCGTCTCTCGCTGGAACACTTCAACAGC
CTCTGGGACCAAAACAGGCCATACTGAAGCACTCTGCTGAAAACACACACTTCAACTCCAGCAGTGAAGCTGAG
CAGAAGACACAGCGCTTCTACCTATCTTGCCCTACAGGTGGAGCGCATTGAGGAGCTGGGAGCCAGCCACTGAGA
GACCTCATTGAGAAGATTGGTGGTTGGAAACATTACGGGGCCCTGGGACAGGACAACCTTATGAGGAGTTGGAAG
GCAGTAGCAGGGACCTACAGGGCCACCCCATCTTCAACGCTCTACATCAGTCCGACTCTTAAGAGTTCCAAACAGC
AATGTTATCCAGGTGGACAGTCTGGGCTCTTTCTGCCCTCTCGGGATTACTACTTAAACAGAACTGCCAATGAG
AAAGTGCTCAGTGCCATCTGGATTACATGGAGGAACCTGGGATGCTGCTGGTGGGCGGCCACCTCCACGAGG
GAGCAGATGCAGCAGGTGCTGGATTGGAGATACAGCTGGCCAACATCACAGTGCCTCAGGACAGCGGGCGGAC
GAGGAGAAGATCTACCAAGATTGAGCATTTCGGAGCTGCAGGCTCTGGCGCCCTCCATGGAGTGGCTTGAGTTT
CTGTCTTTCTTGCTGTCAACATTGGAGTTGAGTGAGCTCTGAGCCTGTGGTGGTGTATGGGATGATTATTTGCAG
CAGGTGTCAAGCTCATCAACCGCACGGAACCAAGCATCCTGAACAATTACCTGATCTGGAACCTGGTGCAAAAG
ACAACCTCAAGCTTGACCGACGCTTTGAGTCTGCACAAGAGAAGCTGCTGGAGACCTCTATGGCACTAAGAAG
TCCTGTGTGCCGAGGTGGCAGACCTGCATCTCCAACACGGATGACGCCCTTGGCTTTGCTTTGGGGTCACTCTTC
GTGAAGGCCACGTTTGACCGGCCAAAGCAAAGAAATTGCAGAGGGGATGATCAGCGAAATCCGGACCGCATTTGAG
GAGGCCCTGGGACAGCTGGTTTGATGGATGAGAAGACCCGCGAGGACCCAGGAGAAAAGCAGATGCCATCTAT
GATATGATTGGTTTCCAGACTTATCTCGGAGCCCAAAGAGCTGGATGATGTTTATGACGGGTACGAAATTTCT
GAAGATTCTTTCTTCCAAAACATGTTGAATTTGTACAACCTCTCTGCCAAGGTTATGGCTGACCACTCCGCAAG
CCTCCAGCCGAGACCAAGTGGAGCATGACCCCCAGACAGTGAATGCCCTACTACCTTCCAACATAAGAAATGAGATC
GTCTTCCCGCTGGCATCTGCAGGCCCTTCTATGCCCGCAACCAACCCAGGCGCTGAACCTCGTGGGCATC
GGTGTGCTCATGGGCCATGAGTTGACGCTGCCCTTTGATGACCAAGGGCGCGAGTATGACAAAGAAGGGAACCTG
CGGCCCTGGTGGCAGAAATGAGTCCCTGGCAGCCTTCCGGAACCAACAGGCCCTGCATGGAGGAACAGTACAATCAA
TACCAGGTCAATGGGGAGAGGCTCAACGGCCGCCAGACGCTGGGGGAGAACATTACTGACAAACGGGGGGCTGAAG
GCTGCTCAATGCTTACAAGCATGGCTGAGAAAAGCATGGGGAGGAGCAACCTGCGAGCCCTGGGGCTCACC
AACCACAGCTCTTCTCGTGGGATTTGCCAGGTGTGGTGTCTCGGTCCGCACACCAGAGAGCTCTCACAGGGGG
CTGGTGACCGACCCCCACAGCCCTGCCGCTTCCGCGTGTCTGGGCACTCTCTCAACTCCCGTGACTTCTGTGCGG
CACTTCCGCTGCCCTGTGCGCTCCCCATGAACCCAGGGCAGCTGTGTGAGGTGTGGT~~TAGAC~~CTGGATCAGGGGA
GAAATGGCCAGCTGTCAACAGACCTGGGCGAGCTCTCCTGACAAAGCTGTTTGTCTTGGGTTGGGAGGAAGCAA
ATGCAAGCTGGGCTGGGCTAGTCCCTCCCCCCCCAGGTGACATGAGTACAGACCTCTCAATCACCACATTTG
TGCTCTGCTTTGGGGGTGCCCTGCTCCAGCAGAGCCCCACACTTCACTGTGACATCTTCCGTGTCAACCT
GCCTGGAAGAGGTCTGGGTGGGGAGGCCAGTTCCCATAGGAAGGAGTCTGCC

109701-5049269

FIGURE 225

MNVALQELGAGSNVGFQKGTRQLLGSRTQLELVLAGASLLLAALLLGCLVALGVQYHRDP SH
STCLTEACIRVAGKILES LDRGVSPCEDFYQFSCGGWIRRNPLPDGRSRWNTFNSLWDQ NQA
ILKHLLENTTFNSSSEAEQKTQRFYLSCLQVERIEELGAQPLRDLIEKIGGWNITGPWDQDN
FMEVLKAVAGTYRATPFFT VYISADSKSSNSNVIQVDQSGFLFSPRDYYLNRTANEKVLTAY
LDYMEELGMLLGGRPTSTREQMQQVLELEIQLANITV PQDQRRDEEKIYHKMSISELQALAP
SMDWLEFLSFLLSPLELSDSEPVVVYGM DYLQQVSELINRTEPSILNNYLIWNLVQKTTSSL
DRRFESAQEKLLETLYGTTKSCVPRWQTCISNTDDALGFALGSLFVKATFDRQSKEIAEGMI
SEIRTA FEEALGQLVWMDEKTRQAAKEKADAIYDMIGFPDFILEPKELDDVYDGYEISED SF
FQNMLNLYNFS AKVMADQLRKPPSRDQWSMTPQTVNAYYLP TKNEIVFPAGILQAPFYARNH
PKALNFGGIGVVMGHELTHAFDDQGREYDKEGNLRPWWQNESLAAFRNHTACMEEQYNQYQV
NGERLNGRQTLGENITDNGGLKAAYNAYKAWLRKHGEEQQLPAVGLTNHQLFFVGFAQVWCS
VRTPESSHEGLVTDPHSPARFRVLGTLSNSRDFLRHFGCPVGSPMNPQG LCEVW

Type II Transmembrane domain:

amino acids 32-57

CCCCGGCCCTCCGGCCCTCCGCACTCCCGGCTCCCTCCCTCCGGCCGGCTCCCGGGCCGCTCCCTCCCTCCCTCCCG
CAGCTGTCCTCGGTCGCTCATCCGACGACCTCCCGGCCCGCCCGGCCCGCTGCTGCTCTCCGGGTCGTGCTGCTCC
CGGCTCCCGGCCCGGCCCGGCCCGCCCGCCAGAGCCCCCGTGTGCCCACTCGTTCTGAGAAGGAGCCGCTGCC
CGTTTCGGGAGAGCGGACAGTAGTAGTGGGCGCCCGGGGAGGCGCGCGGGCGGGAGATCGGGCTCCGGGCGAGTCAAGCGC
CAGCCCGAGAGGGGGCGCGGGCGCGAGTGGCTCGCGCGCGCGCGCGCGAGGGTGGGCGGGGCGAGAGGGGCG
CGGGTCTGCTGGGACCGCGGGACCGCGGGGACGCCCGCGGGGCGGACACAGCGCGAGCTGGGCGAGCGGGCTCCGAG
CAAGCCCGTCCCGGACAGCTGACCTTTCCGGCGAGGCTCTATGCCCTTGGACGAGACAGTGGCACCGGACCTGAGC
GGAGCCATCTCGGGGTGATGCGCTCGCTGTGCGCTTGCAGCGCGCAGTGGGGTCCGCGTATCACAGGGGCCCTGG
CAGGGTCAGTCTGCAAGAACTCAAACCCAGAGTCCCAACCCCGGCTTGTGGGCGAGCGCGCGCGCTGCGGGCACA
CTGTGCGACAGCTGCCCCAGGACTTCGTGGGCGTGTCTGACAGGGCGAGCTGCGAGCGGTGGACAGAGCC
AGTCTCGCTGCTGGCCTCTAGCTCCGCTTCTCTATCTCTCTACGAGCGGCTGAGCGCCCTACAGGATCCGCTCT
CTAGACTCCAAATGGCAGTGTCTCTTTTGAAGACCCCTGACGCCCCACAAGATGGCTGGTCTGTGGGGTGTG
CGGGGACAGTCTCGGTTGTCTTCGCGGCTCTTAGGCGACAAAGAGCTGATGTGCACTTGTGACACTCTGAC
CCCTTCAGGGAGGTTCTGGGGGCTCTCTATCCGGACACGGGCGCTGTCTCCAGAGACTTCACTGACACTCTGAC
TCTTAGAGCCGCCCCACAGCAGGGCTGAGGGGGCATCACCTGTCTCACTCTCACTGACACAGAGGACTCTTGCTGA
TTTTTTGCTGCTCTCCGAGGCGCTTGCAGGACTAACCCAGTTCCCTTGTAGGCTCCAGACTTACACAGGGGCA
GCTATCTGCGAGAACTTCAGGCCAATGTCTCAGGCCAGGAACAGGCTTTGTGAGTGTCTGCCAACCTGACAGT
CCAGGAGATGAGACTGCTGGTGCTGGGGAGAGCTGAGATGCGCTGAGATGGGACGAGCAGAGGCTGCGCAT
CAGTGACACACTTGTCTGCGGAAGAGCTGCGAGCTGCTGCAAGTGTCTTTTGTGGGGCTAATGCCGTGATCC
AGTCTAAACCGGTTGCTCGGCGCTCAGGCCGACTCACTCTGTCTAGGAAGATGGCNCCTGATCTCCAGTGTCAATT
GTGAGGACACAGGTGAGTGTGTGGCGATGACATGGAAGCAAGGCTACAGCGAGGATCAGCCCACTCTGCTCT
GTGCGACATGTGCTGACATCTCTCCCTCTGCCCGAGCGCGTGGTATCTGCGCTGGGCTGGGTTGCGCGAGGGC
TCTATAGTCTGCTGCAAGATGAGCTTTCTCTGAAGCTGGGCAACAGGACTTCCGACCGAGAGCTTCCGGGCA
ACGTGGCTGCCCTGCCCTACTGTGGGGGACGTGCGGCCCTGCGCGTGGCCCTAGCAGGAGCCTTGGTGTCTACC
CGCTGTGAAGAGGCGACAGCAGGCGACGGCCGCTGTTCTTGATACCACTATGCTCACTGTCACTGACATGAAGTGT
CTGTGGCTGGGCTTGGTGGCTCAGAAACAGGCACTGTCTACTGCCACCTCTTGGGCTCTCGGAAGCAGCGGCC
TCGGCGGCTGCTGAAGGAACTTATGGCTCAGAGCGCCAGGGTGTGTGAAGACCTGAGCGCGGAATCTGCTGCG
GCACCTGGGCAAGAGCACTGGCTTCCGTGATGATACCCACAGGTAGTCCGACAGGGAGGATCTCCGAGGGCGCT
CTCTCTCCAGGTGACATATGACCAACCAATGTGAGTTGGTGGCATCTGCGCTGAGCGGCGCGGGCCGAGGGGT
CGGGCGGCTTGGGGCTCCGGATACAGCTCTGTCTGCGCGCTGTGGTGTGCTGGTCTCCCGGCCCTAGCGCCGCG
CAAACTGTGTGTCTTGGGGCGGCCGAGACCCCAACACTGCTTCTTGTAGGGGCGAGCAGCGCCCCACGGGGCT
TGCTGGGCGGCCCACTACGACCCGCTCTGTCTCATCTGTGCATCTGTGCAGAGCAAGCGCTGATCTGCTGCGCGGT
TGGTCTGAGCAACCGCCGACGCTGCCACACCGCTGTGAGGCTCCGACAGGTGCTGCTGTTTTCCTGGGTCGCTA
TTTTGATGTGATACCGGAGCTGGCGGGCAGCGGCTAGCGGTTGGACCCCGTGTGCTGCCCCCTTTGGCTTAATTA
TGTGTCTGTCTGCACCTGCAAGAGGGGGGACCTGAGGAGGTGTCACTGTGAGAAGTGCAGTGTCCCGGCTGGC
CTGTGCCAGGCTGTGCGTGTCAACCCACCGACTCTGTGTCAAAAGCTGTCCAGGTGAGGCCACCCCGAGCTGGG
GGACCTGCTCAGGCTGATGGGCGCCGGGCTGCGCTTGTGTGGGCAATGTGTTCCGAGAGTCAAGCTGAGTGC
CCCTCAGTGCCTCCGTTTGAAGAGATGACCTGTATCACCTGCAAGATGTGGGTAACTGGGAGAGCAGGCTTGT
CTGAGGTGGGTATCTGGGAGCTCTGTGCTGTGAGTAGGAGAACCTTCCGAGGAGGTCCCTGAAGAAGCTGAAGTCT
TGTGTCTCCAGTGCCTTGGGGGACACTCAGTGTGTGCTCTGTCTGTGACAGGCGAGGGTGTCTACTGTGAGC
GGGATGACTGTGTACCTGCCATCTGTCTCTGTGGCTCGGGGAGGAGAGTCTGATGTCTTCCGCTGCACGCGTACC
GGCGGCTGATGTAGGAGGAGTCCAGGTTCAGCAGTCTGAGTGTGAGGGCTCACTGCTGTGGACTCTGTGATCAG
GGAAGGGGACACTCACTGTGTGACGAGAACAGTGCAGCTGCTCTCAAGTGTGCATTCCAATCACCTCCCTCAGCA
ACTCGTGGAAATTTGTTATTTATGACTTTTCTTTTCAAAAGATTTTGAAGTCTGAAGCTCAGAAATTAAGCAACAG
ATGAAGTCAACCCAGCTGTGTGCTGACCTGCTTTTGAAGAAATCTGGCTTTCTGGGACCAAGGCGAGGATGCTT
TGAGCTGCCCTCTATGCTCTCTGTGCTCTCTCACTCTCCCTCTCCCAACTCTCTCTCTCTGTCTCT
AGCAGGCCACAGAGACCGAGACTGATCCAGAGCTGTGAGAAAGAGAGCCGGAAGCTCTTAGGGAGCAGCAGAGGGCC
AAGTGAACCAAGAGATGGGCTGAGCTGGGGAAGGGTGCCATCGAGGACTTTCGATCTCTGTGGGAGAG
CCAGTGCTCTTTGCTCTCTGTCTGCTCTACTCTACCCACCACCTACTCTTGGGAACACAGCTCCAACAGGG
GAGAGCGAGCTGGGCGACGCGAGGTCAAGCACTCAAGTCTGCCCCGTCCGCTCCGCTCTGTCTCTGAGG
GCCCCACGCTTTCTCTCTGATCAATAATGTCTACTGTGCTGTGGGATTTTAAATTTACTTCACTCAGCAACAG
GGCCCGGACACTCCACTCTCTGCTGCCCTGAGCTGAGCAGAGTCAATTATGAGAGATTTTGTATTTAAAC
ATTTCTTTTTCAGTCTTTGGGACATGAGTTGTGCTTTTGTGGGCAAGCACTGAGTGGGCTGTGTGAGAGGG
CGNAGAGTAGTAGGAGTGTAGAGAGAGGAGCTCTGACATCTTGGGAGCTGAAAGAGACCTGGAGAGGCAGAGGATAG
CTGGCCNNNTTGGCTGGCAATNCCTGGGTTCCGAGAGGGGCTGGGATGTTCTTGAAGTGGTTCAGAGCTCAAG
AATTTGAGAGGAGTAGAAGCAGGATTTTGAATCAAGTTTAGTTTCCCACTGCTGCGCTGTTTGTGACTCTTGT
TTTGAAGTGTCTCAGAGAGAGAACTCAAAGTGTGACACAGCCCTCTCTCCCTCTTCCCTTCCCTTCCCTTTCT
TGACCTTCCCTCCCTCCCTCCCTCCCTCCCTCCCT

FIGURE 227

GGCCGAGCGGGGTGCTGCGCGGCGGCCGTGATGGCTGGTGACGGCGGGGCCGGGCAGGGGA
CCGGGGCCGCGGGCCCGGAGCGGGCCAGCTGCCGGGAGCCCTGAATCACCGCCTGGCCCCGAC
TCCACCATGAACGTCGCGCTGCAGGAGCTGGGAGCTGGCAGCAACGTGGGATTCCAGAAGGG
GACAAGACAGCTGTTAGGCTCACGCACGCAGCTGGAGCTGGTCTTAGCAGGTGCCTCTCTAC
TGCTGGCTGCACTGCTTCTGGGCTGCCTTGTGGCCCTAGGGGTCCAGTACCACAGAGACCCA
TCCCACAGCACCTGCCTTACAGAGGCTGCATTCGAGTGGCTGGAAAAATCCTGGAGTCCCT
GGACCGAGGGGTGAGCCCCCTGTGAGGACTTTTACCAGTTCTCCTGTGGGGGCTGGATT CGGA
GGAACCCCTGCCCCGATGGGCGTTCTCGCTGGAACACCTTCAACAGCCTCTGGGACCAAAC
CAGGCCATACTGAAGCACCTGCTTGAAAAACCCACCTTCAACTCCAGCAGTGAAGCTGAGCA
GAAGACACAGCGCTTCTACCTATCTTGCTACAGGTGGAGCGCATTGAGGAGCTGGGAGCCC
AGCCACTGAGAGACCTCATTGAGAAGATTGGTGGTTGGAACATTACGGGGCCCTGGGACCAG
GACAACTTTATGGAGGTGTTGAAGGCAGTAGCAGGGACCTACAGGGCCACCCCATTTCTCAC
CGTCTACATCAGTGCCGACTCTAAGAGTTCCAACAGCAATGTTATCCAGGTGGACCAGTCTG
GGCTCTTTCTGCCCTCTCGGGATTACTACTTAAACAGAACTGCCAATGAGAAAGTAAGGAAC
ATCTTCCGAACCCCCATCCCTACCCCTGGCTGAGCTGGGCTGATCCCTGTTGACTTTTCCCT
TTGCCAAGGGTCAGAGCAGGGAAGGTGAGCCTATCCTGTCACCTAGTGAACAAACTGCCCTT
CCTTTCTTTCTTCTTTTCTTCTCCTCCCTCCCTCCCTTTCTTCCCTTTTCTTCTTCTTCC
TCTTATTCTTCTAGTAGGTTTCATAGACACCTACTGTGTGCCAGGTCCAGTGGGGGAATTG
GAGATATAAGTTTCCGAGCCATTGCCACAGGAAGCGTTCAGTGTGATGGGTTCATGGACCT
AGATAGGCTGATAACAAAGCTCACAAGAGGGTCCTGAGGATTCAGGAGAGACTTATGGAGCC
AGCAAAGTCTTCTGAAGAGATTGCATTTGAGCCAGGTCCGTAG

FIGURE 228

ATGCCTACTACCTTCCAAC TAAGAATGAGATCGTCTTCCCCGCTGGCATCCTGCAGGCCCC
TTCTATGCCCGCAACCACCCCAAGGCCCTGAAC TTCCGGTGGCATCGGTGTGGTCATGGGCCA
TGAGTTGACGCATGCC TTTGATGACCAAGGGCGCGAGTATGACAAAGAAGGGAACCTGCGGC
CCTGGTGGCAGAATGAGTCCCTGGCAGCCTTCCGGAACCAACGGCCTGCATGGAGGAACAG
TACAATCAATACCAGGTCAATGGGGAGAGGCTCAACGGCCGCCAGACGCTGGGGGAGAACAT
TGCTGACAACGGGGGGCTGAAGGCTGCC TACAATGCTTACAAAGCATGGCTGAGAAAGCATG
GGGAGGAGCAGCAACTGCCAGCCGTGGGGCTCACCAACCACCAGCTCTTCTTCGTGGGATT
GCCCAGGTGTGGTGCTCGGTCCGCACACCAGAGAGCTCTCACGAGGGGCTGGTGACCGACCC
CCACAGCCCTGCCCCGTTCGCGTGCTGGGCAC TCTCTCCAAC TCCCGTGA CTCTCTGCGGC
ACTTCGGCTGCCCTGTCTGGCTCCCCATGAACCCAGGGCAGCTGTGTGAGGTGTGGTAGACC
TGGATCAGGGGAGAAATGGCCAGCTGTCAACAGACCTGGGGCAGCTCTCTGACAAAGCTGT
TTGCTCTTGGGTTGGGAGGAAGCAAATGCAAGCTGGGCTGGGTCTAGTCCCTCCCCCCCACA
GGTGACATGAGTACAGACCTCTCTCAATCACCACATTGTGCCTCTGCTTTGGGGGTGCCCT
GCCTCCAGCAGAGCCCCCACCATTCACTGTGACATCTTTCCGTGTACCTGCGCTGGAAGAG
GTCTGGGTGGGGAGGCCAGTTCCCATAGGAAGGAGTCTGCCTCTTCTGTCCCCAGGCTCACT
CAGCCTGGCGGCCATGGGGCCTGCCGTGCCTGCCCCACTGTGACCCACAGGCCTGGGTGGTG
TACCTCCTGGACTTCTCCCCAGGCTCACTCAGTGC GCACCTTAGGGGTGGACTCAGCTCTGT
TGGCTCACCTTCACGGGCTACCCCCACCTCACCTGTGCTCCTTGTGCCACTGCTCCAGTG
CTGCTGTGACCTTCACTGACAGCTCCTAGTGGAAGCCCAAGGCCTCTGAAAGCCTCCTGC
TGCCCACTGTTTCCCTGGGCTGAGAGGGGAAGTGCATATGTGTAGCGGGTACTGGTTCCTGT
GTCTTAGGGCACAAGCCTTAGCAAATGATTGATTCTCCCTGGACAAAGCAGGAAAGCAGATA
GAGCAGGGAAAAGGAAGAACAGAGTTTATTTTACAGAAAAGAGGGTGGGAGGGTGTGGTCT
TGGCCCTTATAGGACC

FIGURE 229

CCCACGCGTCCGAGCGCGCCGAGAATTAGACACACTCCGGACGCGGCCAAAAGCAACCGAGA
GGAGGGGAGGCAAAAACACCGAAAAACAAAAAGAGAGAAACAAACCCCAACACTGGGGGTGG
GGGGAAGAAAGAAAGAAAGAAACCCACCCACCACCAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAATCCTGTGGCGCGCCGCTGGTTCCTCCGGGAAGACTCGCCAGCACCCAGGGG
TGGGGGAGTGCAGCTGAAAGCTGCTGGAGAGTGAGCAGCCCTAGCAGGATGGACATGATG
CTGTTGGTGCAGGGTGCTTGTGTCTCGAACCAGTGGCTGGCGCGGTGCTCCTCAGCCTGTG
CTGCCTGTCTACCTCTGCCTCCCGGCTGGACAGAGTGTGGACTTCCCTTGGCGGGCCGTGG
ACAACATGATGGT CAGAAAAGGGGACACGGCGGTGCTTAGGTGTTATTTGGAAGATGGAGCT
TCAAAGGGTGCTGGCTGAACCGGTCAAGTATATTTTTTGGCGGAGGTGATAAGTGGTCAGT
GGATCCTCGAGTTTCAATTTCAACATTGAATAAAAGGGACTACAGCCTCCAGATACAGAATG
TAGATGTGACAGATGATGGCCCATACAGTGTCTGTTCAGACTCAACATACACCCAGAACA
ATGCAGGTGCATCTAACTGTGCAAGTTCTCTCCTAAGATATATGACATCTCAAATGATATGAC
CGTCAATGAAGGAACCAACGTCACTCTTACTTGTGTGGCCACTGGGAAACCAGAGCCTTCCA
TTTTCTGGCGACACATCTCCCCATCAGCAAAACCAATTGAAAATGGACAATATTTGGACATT
TATGGAATTACAAGGGACCGGCTGGGGAATATGAATGCAGTGCAGGAAATGCTGTGTCAATT
CCCAGATGTGAGGAAAGTAAAGTTGTTGTCAACTTTGCTCCTACTATT CAGGAAATTAAT
CTGGCACCGTGACCCCCGGACGCAAGTGGCCTGATAAGATGTGAAGGTGCAGGTGTGCGCGCT
CCAGCCTTTGAATGGTACAAAGGAGAGAAGAAGCTCTTCAATGGCCAACAAGGAATTATTAT
TCAAATTTTAGCACAAAGATCCATTCTCACTGTACCAACGTGACACAGGAGCACTTCGCA
ATTATACCTGTGTGGCTGCCAACAAGCTAGGCACAACCAATGCGAGCCTGCCTCTTAACCT
CCAAGTACAGCCAGTATGGAATTACCGGGAGCGCTGATGTTCTTTTCTCCTGCTGGTACCT
TGTGTTGACCTGTCTCTCTTACCAGCATACTCTACCTGAAGATGCCATTCTCAATATAA
TTCAAAGACCCATAAAAGGCTTTTAAAGGATTCTCTGAAAGTGCTGATGGCTGGATCCAATCT
GGTACAGTTTGTATAAGCAGCGTGGGATATAATCAGCAGTGTCTACATGGGGATGATCGCC
TTCTGTAGAATTGCTCATTTATGTAAATACCTTTAATTCTACTCTTTTTTATTAGTACATTA
CCTTGTAAGCAGTACACATTGTCTTTTTTTTAAAGCGTGAAGCTCTGAAATTACTTTTAG
AGGATATTAATTGTGATTTTCATGTTTGTAAATCTACAACTTTTCAAAGCATT CAGTCATGGT
CTGTAGGTTGCAGGCTGTAGTTTACAAAAACGAATATTGCAGTGAATATGTGATTCTTTAA
GGCTGCAATACAAGCATT CAGTTCCTGTTTCAATAAGAGTCAATCCCAATTTACAAAGATG
CATTTTTTTCTTTTTTGATAAAAAAGCAAATAATATTGCCTTCAGATTATTTCTTCAAATA
TAACACATATCTAGATTTTTCTGCTTGCATGATATTCAGGTTTCAGGAATGAGCCTTGTAA
ATAACTGGCTGTGCAGCTCTGCTTCTCTTCCCTGTAAGTTGCAGATGGGTGCTCCTTCATAC
AATAATATTTTCTCTTGTCTCCAACTAATATAAAATGTTTTGCTAAATCTTACAATTTGA
AAGTAAAAATAAACAGAGTGATCAAGTTAAACCATACACTATCTCTAAGTAACGAAGGAGC
TATTCGACTGTAAAAATCTCTTCTGCACTGACAATGGGGTTTGAGAATTTTGGCCACACT
AACTCAGTTCCTGTGATGAGAGACAATTTAATAACAGTATAGTAAATATACCATATGATTTT
TTTAGTTGTAGCTAAATGTTAGATCCACCGTGGGAAATCATCTCCCTTTAAAAATGACAGCACA
GTCCACTCAAAGGATTGCGCTAGCAATACAGCATCTTTCTTCTACTAGTCCAAGCCAAAA
TTTTAAGATGATTTGT CAGAAAGGGCACAAAGTCTATCACCTAATATTACAAGAGTTGGTA
AGCGCTCATATTAATTTTATTTTGTGGCAGGTATTATGACAGTTCGACCTGGAGGGTATGGA
TATGGATATGGACGTTCCAGAGACTATAATGGCAGAAACAGGGTGGTTATGACCGCTACTC
AGGAGGAAATTACAGAGACAAATTATGACAACCTGAAATGACAGATGCACATAATATAGATACA
CAAGGAATAATTTCTGATCCAGGATCGTCCTTCCAAATGGCTGTATTTATAAGGTTTTTGG
AGCTGCAGTGAAGCATCTTATTTTATAGTATATCAACCTTTTGTTTTTTAAATTGACCTGCCA
AGGTAGCTGAAGACCTTTTATAGACAGTTCCATCTTTTTTTTAAATTTTCTGCCTATTTAA
AGACAATTTATGGGACGTTTGTCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 230

MMLLVQGACCSNQWLAAVLLSLCCLLPSCLPAGQSVDFPWAAVDNMMVRKGD TAVLR CYLED
GASKGAWLNRSSIIIFAGGDKWSVDP RVSISTLNKR DYSLQIQNV DVTDDGPYTCSVQTQHTP
RTMQVHLTVQVPPKIYDISNDMTVNEGTNVTLTCLATGKPEPSISWRHISPSAKPFENGQYL
DIYGITRDQAGEYEC SAENAVSFDPVRKV KVVVNFAPTIQEIKSGTVTPGRSGLIRCEGAGV
PPPAFEWYKGEKKLFNGQQGII IQNFSTRSILTVTNVTQEHFGNYTCVAANKLGT TNASLPL
NPPSTAQYGITGSADVLFSCWYLVLTLS SFTSIFYLKNAILQ

Important features of the protein:

Signal peptide:

amino acids 1-31

Transmembrane domain:

amino acids 326-345

N-glycosylation sites.

amino acids 71-75, 153-157, 273-277, 284-288, 292-296, 305-309

Casein kinase II phosphorylation site.

amino acids 147-151, 208-212, 224-228

Tyrosine kinase phosphorylation site.

amino acids 178-186

N-myristoylation sites.

amino acids 7-13, 63-70, 67-73, 151-157, 239-245, 291-297,
302-308, 319-325

Myelin P0 protein:

amino acids 92-121

FIGURE 231

AGTGGTTCGATGGGAAGGATCTTCTCCAAGTGGTTCCTCTTGAGGGGAGCATTCTGCTGG
CTCCAGGACTTTTGGCCATCTATAAAGCTTGGCAATGAGAGAAATAAGAAAATTCTCAAGGAGGA
CGAGCTCTTGAGTGAGACCCAAACAAGCTGCTTTTCCACAAATTGCAATGGAGCCTTTTCGAAA
TCAATGTTCCAAAGCCCAAGAGGAGAAATGGGGTGAACTTCTCCCTAGCTGTGGTGGTCATC
TACCTGATCCTGCTCACCGCTGGCGCTGGGCTGCTGGTGGTCCAAGTTCTGAATCTGCAGGC
GCGGCTCCGGGTCTTGAGATGTATTTCTCAATGACACTCTGGCGGCTGAGGACAGCCCCGT
CCTTCTCCTTGCTGCAGTCAGCACACCTTGGAGAACACTGGCTCAGGGTGCATCGAGGCTG
CAAGTCTGCAGGCCAACTCACCTGGGTCCGCGTCAGCCATGAGCACTTGCTGCAGCGGGT
AGACAACCTCACTCAGAACCCAGGGATGTTCAGAATCAAAGGTGAACAAGGCGCCCCAGGTC
TTCAAGGTCAACAAGGGGGCCATGGGCATGCTGGTGCCCTGGCCCCCGGGACCACCTGCT
GAGAAGGGAGCCAAGGGGGCTATGGGACGAGATGGAGCAACAGGCCCTCGGGAGCCCCAAGG
CCCACCGGGAGTCAAGGGAGAGGCGGGCTCCAAGGACCCAGGGTGCTCCAGGGAAGCAAG
GAGCCACTGGCACCCCCAGGACCCCCAAGGAGAGAAGGGCAGCAAAGGCGATGGGGGTCTCATT
GGCCCCAAAAGGGGAAACTGGAACTAAGGAGAGAAAAGGAGACCTGGGTCTCCAGGAAGCAA
AGGGGACAGGGGCATGAAAGGAGATGCAGGGGTCTATGGGGCTCTTGAGGCCAGGGGAGTA
AAGGTGACTTCGGGAGGCCAGGCCACCAGGTTTGGCTGGTTTTTCTTGAGCTAAAGGAGAT
CAAGGACAACCTGGACTGCAGGGTGTTCGGGCCCTCTGCTGAGTGGGACACCAGGTGC
CAAGGTGAGCCTGGCAGTGTGCTGCTCCCTGGGCGAGCAGGACTTCAGGGAGCCCCGGGA
GTCCAGGAGCCACAGGCCTGAAAGGAAGCAAAGGGGACACAGGACTTCAAGGACAGCAAGGA
AGAAAAGGAGAAATCAGGAGTTCAGGCCCTGCAGGTGTGAAGGGAGAACAGGGGAGCCCAGG
GCTGGCAGGTCCCAAGGGAGCCCTGGACAAGCTGGCCAGAAGGGAGACCAGGGAGTGAAG
GATCTTCTGGGGAGCAAGGAGTAAAGGGAGAAAAGGTGAAAGAGGTGAAAACCTCAGTGTCC
GTCAGGATTGTGCGCAGTAGTAACCGAGGCCGGGCTGAAGTTTACTACAGTGGTACCTGGGG
GACAAATTGCGATGACGAGTGGCAAAATTCTGATGCCATTGTCTTCTGCCGCATGCTGGGTT
ACTCCAAAGGAAGGGCCCTGTACAAAGTGGGAGCTGGCACTGGGCAGATCTGGCTGGATAAT
GTTCACTGTGCGGGCACGGAGAGTACCTGTGGAGCTGCACCAAGAATAGCTGGGGCCATCA
TGACTGCAGCCACGAGGAGGACGCAGGCGTGGAGTGCAGCGTCTGACCCGGAACCCCTTCA
CTTCTCTGCTCCCAGGTGTCTCGGGCTCATATGTGGGAAGGCAGAGGATCTCTGAGGAGT
TCCCTGGGGACAACCTGAGCAGCCTCTGGAGAGGGGCCATTAATAAAGCTCAACATCATTGA

Figure 1. The effect of the concentration of the *Agaricus bisporus* spores on the growth of *Agaricus bisporus* on the substrate. The concentration of the spores was 10⁴ spores/g (a), 10⁵ spores/g (b), 10⁶ spores/g (c), 10⁷ spores/g (d), 10⁸ spores/g (e), 10⁹ spores/g (f), 10¹⁰ spores/g (g), 10¹¹ spores/g (h), 10¹² spores/g (i), 10¹³ spores/g (j), 10¹⁴ spores/g (k), 10¹⁵ spores/g (l), 10¹⁶ spores/g (m), 10¹⁷ spores/g (n), 10¹⁸ spores/g (o), 10¹⁹ spores/g (p), 10²⁰ spores/g (q), 10²¹ spores/g (r), 10²² spores/g (s), 10²³ spores/g (t), 10²⁴ spores/g (u), 10²⁵ spores/g (v), 10²⁶ spores/g (w), 10²⁷ spores/g (x), 10²⁸ spores/g (y), 10²⁹ spores/g (z), 10³⁰ spores/g (aa), 10³¹ spores/g (ab), 10³² spores/g (ac), 10³³ spores/g (ad), 10³⁴ spores/g (ae), 10³⁵ spores/g (af), 10³⁶ spores/g (ag), 10³⁷ spores/g (ah), 10³⁸ spores/g (ai), 10³⁹ spores/g (aj), 10⁴⁰ spores/g (ak), 10⁴¹ spores/g (al), 10⁴² spores/g (am), 10⁴³ spores/g (an), 10⁴⁴ spores/g (ao), 10⁴⁵ spores/g (ap), 10⁴⁶ spores/g (aq), 10⁴⁷ spores/g (ar), 10⁴⁸ spores/g (as), 10⁴⁹ spores/g (at), 10⁵⁰ spores/g (au), 10⁵¹ spores/g (av), 10⁵² spores/g (aw), 10⁵³ spores/g (ax), 10⁵⁴ spores/g (ay), 10⁵⁵ spores/g (az), 10⁵⁶ spores/g (ba), 10⁵⁷ spores/g (bb), 10⁵⁸ spores/g (bc), 10⁵⁹ spores/g (bd), 10⁶⁰ spores/g (be), 10⁶¹ spores/g (bf), 10⁶² spores/g (bg), 10⁶³ spores/g (bh), 10⁶⁴ spores/g (bi), 10⁶⁵ spores/g (bj), 10⁶⁶ spores/g (bk), 10⁶⁷ spores/g (bl), 10⁶⁸ spores/g (bm), 10⁶⁹ spores/g (bn), 10⁷⁰ spores/g (bo), 10⁷¹ spores/g (bp), 10⁷² spores/g (bq), 10⁷³ spores/g (br), 10⁷⁴ spores/g (bs), 10⁷⁵ spores/g (bt), 10⁷⁶ spores/g (bu), 10⁷⁷ spores/g (bv), 10⁷⁸ spores/g (bw), 10⁷⁹ spores/g (bx), 10⁸⁰ spores/g (by), 10⁸¹ spores/g (bz), 10⁸² spores/g (ca), 10⁸³ spores/g (cb), 10⁸⁴ spores/g (cc), 10⁸⁵ spores/g (cd), 10⁸⁶ spores/g (ce), 10⁸⁷ spores/g (cf), 10⁸⁸ spores/g (cg), 10⁸⁹ spores/g (ch), 10⁹⁰ spores/g (ci), 10⁹¹ spores/g (cj), 10⁹² spores/g (ck), 10⁹³ spores/g (cl), 10⁹⁴ spores/g (cm), 10⁹⁵ spores/g (cn), 10⁹⁶ spores/g (co), 10⁹⁷ spores/g (cp), 10⁹⁸ spores/g (cq), 10⁹⁹ spores/g (cr), 10¹⁰⁰ spores/g (cs), 10¹⁰¹ spores/g (ct), 10¹⁰² spores/g (cu), 10¹⁰³ spores/g (cv), 10¹⁰⁴ spores/g (cw), 10¹⁰⁵ spores/g (cx), 10¹⁰⁶ spores/g (cy), 10¹⁰⁷ spores/g (cz), 10¹⁰⁸ spores/g (da), 10¹⁰⁹ spores/g (db), 10¹¹⁰ spores/g (dc), 10¹¹¹ spores/g (dd), 10¹¹² spores/g (de), 10¹¹³ spores/g (df), 10¹¹⁴ spores/g (dg), 10¹¹⁵ spores/g (dh), 10¹¹⁶ spores/g (di), 10¹¹⁷ spores/g (dj), 10¹¹⁸ spores/g (dk), 10¹¹⁹ spores/g (dl), 10¹²⁰ spores/g (dm), 10¹²¹ spores/g (dn), 10¹²² spores/g (do), 10¹²³ spores/g (dp), 10¹²⁴ spores/g (dq), 10¹²⁵ spores/g (dr), 10¹²⁶ spores/g (ds), 10¹²⁷ spores/g (dt), 10¹²⁸ spores/g (du), 10¹²⁹ spores/g (dv), 10¹³⁰ spores/g (dw), 10¹³¹ spores/g (dx), 10¹³² spores/g (dy), 10¹³³ spores/g (dz), 10¹³⁴ spores/g (ea), 10¹³⁵ spores/g (eb), 10¹³⁶ spores/g (ec), 10¹³⁷ spores/g (ed), 10¹³⁸ spores/g (ee), 10¹³⁹ spores/g (ef), 10¹⁴⁰ spores/g (eg), 10¹⁴¹ spores/g (eh), 10¹⁴² spores/g (ei), 10¹⁴³ spores/g (ej), 10¹⁴⁴ spores/g (ek), 10¹⁴⁵ spores/g (el), 10¹⁴⁶ spores/g (em), 10¹⁴⁷ spores/g (en), 10¹⁴⁸ spores/g (eo), 10¹⁴⁹ spores/g (ep), 10¹⁵⁰ spores/g (eq), 10¹⁵¹ spores/g (er), 10¹⁵² spores/g (es), 10¹⁵³ spores/g (et), 10¹⁵⁴ spores/g (eu), 10¹⁵⁵ spores/g (ev), 10¹⁵⁶ spores/g (ew), 10¹⁵⁷ spores/g (ex), 10¹⁵⁸ spores/g (ey), 10¹⁵⁹ spores/g (ez), 10¹⁶⁰ spores/g (fa), 10¹⁶¹ spores/g (fb), 10¹⁶² spores/g (fc), 10¹⁶³ spores/g (fd), 10¹⁶⁴ spores/g (fe), 10¹⁶⁵ spores/g (ff), 10¹⁶⁶ spores/g (fg), 10¹⁶⁷ spores/g (fh), 10¹⁶⁸ spores/g (fi), 10¹⁶⁹ spores/g (fj), 10¹⁷⁰ spores/g (fk), 10¹⁷¹ spores/g (fl), 10¹⁷² spores/g (fm), 10¹⁷³ spores/g (fn), 10¹⁷⁴ spores/g (fo), 10¹⁷⁵ spores/g (fp), 10¹⁷⁶ spores/g (fq), 10¹⁷⁷ spores/g (fr), 10¹⁷⁸ spores/g (fs), 10¹⁷⁹ spores/g (ft), 10¹⁸⁰ spores/g (fu), 10¹⁸¹ spores/g (fv), 10¹⁸² spores/g (fw), 10¹⁸³ spores/g (fx), 10¹⁸⁴ spores/g (fy), 10¹⁸⁵ spores/g (fz), 10¹⁸⁶ spores/g (ga), 10¹⁸⁷ spores/g (gb), 10¹⁸⁸ spores/g (gc), 10¹⁸⁹ spores/g (gd), 10¹⁹⁰ spores/g (ge), 10¹⁹¹ spores/g (gf), 10¹⁹² spores/g (gg), 10¹⁹³ spores/g (gh), 10¹⁹⁴ spores/g (gi), 10¹⁹⁵ spores/g (gj), 10¹⁹⁶ spores/g (gk), 10¹⁹⁷ spores/g (gl), 10¹⁹⁸ spores/g (gm), 10¹⁹⁹ spores/g (gn), 10²⁰⁰ spores/g (go), 10²⁰¹ spores/g (gp), 10²⁰² spores/g (gq), 10²⁰³ spores/g (gr), 10²⁰⁴ spores/g (gs), 10²⁰⁵ spores/g (gt), 10²⁰⁶ spores/g (gu), 10²⁰⁷ spores/g (gv), 10²⁰⁸ spores/g (gw), 10²⁰⁹ spores/g (gx), 10²¹⁰ spores/g (gy), 10²¹¹ spores/g (gz), 10²¹² spores/g (ha), 10²¹³ spores/g (hb), 10²¹⁴ spores/g (hc), 10²¹⁵ spores/g (hd), 10²¹⁶ spores/g (he), 10²¹⁷ spores/g (hf), 10²¹⁸ spores/g (hg), 10²¹⁹ spores/g (hh), 10²²⁰ spores/g (hi), 10²²¹ spores/g (hj), 10²²² spores/g (hk), 10²²³ spores/g (hl), 10²²⁴ spores/g (hm), 10²²⁵ spores/g (hn), 10²²⁶ spores/g (ho), 10²²⁷ spores/g (hp), 10²²⁸ spores/g (hq), 10²²⁹ spores/g (hr), 10²³⁰ spores/g (hs), 10²³¹ spores/g (ht), 10²³² spores/g (hu), 10²³³ spores/g (hv

><MW: 52658, pI: 9.16, NX(S/T): 3

amino acids 47-66 (type II)

amino acids 43-47, 83-87, 136-140

amino acids 432-440

amino acids 41-47, 178-184, 253-259, 274-280, 340-346, 346-352,
400-406, 441-447, 475-481, 490-496, 515-521

amino acids 360-364

amino acids 56-78

amino acids 422-471, 488-519

amino acids 151-184, 301-334, 316-349

FIGURE 233

CCCACGCGTCCGAAGGCAGACAAAGGTTCAATTTGTAAGAAGCTCCTTCCAGCACCTCCTCT
CTTCTCCTTTTGGCCAAACTCACCCAGTGAGTGTGAGCATTTAAGAAGCATCCTCTGCCAAG
ACCAAAGGAAAGAAGAAAAAGGGCCAAAAGCCAAAATGAAACTGATGGTACTTGTTTTTCAC
CATTGGGCTAACTTTGCTGCTAGGAGTTCAAGCCATGCCTGCAAATCGCCTCTCTTGCTACA
GAAAGATACTAAAAGATCACAACTGTCACAACCTTCCGGAAGGAGTAGCTGACCTGACACAG
ATTGATGTCAATGTCCAGGATCATTCTGGGATGGGAAGGGATGTGAGATGATCTGTTACTG
CAACTTCAGCGAATTGCTCTGCTGCCCAAAGACGTTTTCTTTGGACCAAAGATCTCTTTCG
TGATTCCTTGCAACAATCAATTGAGAATCTTCATGTATTCTGGAGAACACCATTCTGATTTC
CCACAACTGCACTACATCAGTATAACTGCATTTCAGTTTCTATATAGTGCAATAGAGCAT
AGATTCTATAAATTCTTACTTGTCTAAGACAAGTAAATCTGTGTAAACAAGTAGTAATAAA
AGTTAATTCAATCTAAAAAAAAAAAAA

FIGURE 234

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA52758

<subunit 1 of 1, 98 aa, 1 stop

<MW: 11081, pI: 6.68, NX(S/T): 1

MKLMVLVFTIGLTLLLGVOAMPANRLSCYRKILKDHNCHNLPEGVADLTQIDVNVQDHFWDG
KGCEMICYCNFSELLCCPKDVFFGPKISFVIPCNNQ

Important features:

Signal peptide:

amino acids 1-20

N-glycosylation site.

amino acids 72-76

Tyrosine kinase phosphorylation site.

amino acids 63-71

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 235

CCCACGCGTCCGCGGACGCGTGGGCTGGACCCAGGTCTGGAGCGAATTCAGCCTGCAGGG
CTGATAAGCGAGGCATTAGTGTGATTGAGAGAGACTTTACCCGCGGTGGTGGTTGGAGGGC
GCGCAGTAGAGCAGCAGCACAGGCGCGGGTCCCGGGAGGCGGGCTCTGCTCGCGCCGAGATG
TGGAAATCTCCTTACGAAACGACTCGGCTGTGGCCACCGCGCGCGCCCGCGCTGGCTGTG
CGCTGGGGCGCTGGTGTGGCGGGTGGCTTCTTCTCCTCGGCTTCTCTTCCGGTGGTTTA
TAAAAATCCTCCAATGAAGCTACTAACATTACTCCAAAGCATAATATGAAAGCATTTTTGGAT
GAATTGAAAGCTGAGAACATCAAGAAGTTCCTACATAATTTTACACAGATACCACATTTAGC
AGGAACAGAACAAAACCTTTCAGCTTTCGAAAGCAAATTCATCCCAGTGGAAAGAATTTGGCC
TGGATTCTGTGTGAGCTAGCTCATTATGATGTCTGTGTGCTTACCCAAATAAGACTCATCCC
AACTACATCTCAATAATTAATGAAGATGGAAATGAGATTTTCAACACATCATTATTTGAACC
ACCTCCTCCAGGATATGAAAATGTTTCGGATATTGTACCACCTTTCAGTGCTTCTCTCCTC
AAGGAATGCCAGAGGGCGATCTAGTGTATGTTAACTATGCACGAACTGAAGACTTCTTTAA
TTGGAACGGGACATGAAAATCAATTGCTCTGGGAAAATTTGTAATTGCCAGATATGGGAAAGT
TTTCAGAGGAAATAAGGTTAAAAATGCCAGCTGGCAGGGGCGAAAGGAGTCATTCTCTACT
CCGACCCTGTCTGACTACTTTTGTCTCCTGGGGTGAAGTCTTATCCAGACGGTTGGAATCTTCT
GGAGGTGGTGTCCAGCGTGGAAATATCCTAAATCTGAATGGTGCAGGAGACCCTCTCACACC
AGGTTACCCAGCAAATGAATATGCTTATAGGCGTGGAATTGCAGAGGCTGTTGGTCTTCCAA
GTATTCCTGTTCATCCAATTGGATACTATGATGCACAGAAGCTCCTAGAAAAAATGGGTGGC
TCAGCACCCACAGATAGCAGCTGGAGAGGAAGTCTCAAAGTCCCTACAATGTTGGACCTGG
CTTTACTGGAAACCTTTCTACACAAAAGTCAAGATGCACATCCACTTACCAATGAAGTGA
CGAGAATTTACAATGTGATAGGTACTCTCAGAGGAGCAGTGGAACCAGACAGATATGTCATT
CTGGGAGGTCAACCGGACTCATGGGTGTTTGGTGGTATTGACCCTCAGAGTGGAGCAGCTGT
TGTTTCATGAAATTGTGAGGAGCTTTTGAACACTGAAAAGGAAGGGTGGAGACCTAGAAGAA
CAATTTTGTGTGCAAGCTGGGATGCAGAAGAAATTTGGTCTTCTTGGTCTACTGAGTGGGCA
GAGGAGAATTCAGACTCCTTCAAGAGCGTGGCGTGGCTTATATTAATGCTGACTCATCTAT
AGAAGGAAACTACACTCTGAGAGTTGATTGTACACCCTGATGTACAGCTTGGTACACAACC
TAACAAAAGAGCTGAAAAGCCCTGATGAAGGCTTTGAAGGCAAAATCTCTTTATGAAAGTTGG
ACTAAAAAAAGTCTTCCCCAGAGTTCAGTGGCATGCCAGGATAAGCAAATTTGGGATCTGG
AAATGATTTTGGAGGTGTTCTTCCAACGACTTGAATTTGCTTCAGGCAGAGCAGGGTATCTA
AAAATTTGGGAAACAAACAAATTCAGCGGTATCCACTGTATCACAGTGTCTATGAAACATAT
GAGTTGGTGGAAAAGTTTTATGATCCAATGTTTAAATATCACCTCACTGTGGCCAGGTTTCG
AGGAGGGATGGTGTTTGTAGCTAGCCAAATCCATAGTGTCTCCCTTTTGATTGTGAGATTATG
CTGTAGTTTTTAAGAAAGTATGCTGACAAAATCTACAGTATTTCTATGAAACATCCACAGGAA
ATGAAGACATACAGTGTATCATTGATTCACTTTTTTCTGCAGTAAAGAATTTTACAGAAAT
TGCTTCCAAGTTTCAGTGAGAGACTCCAGGACTTTGACAAAAGCAACCCAATAGTATTAAGAA
TGATGAATGATCAACTCATGTTTCTGGAAAGAGCATTATTGATCCATTAGGGTTACCAGAC
AGGCGCTTTTATAGGCATGTCTATGCTCCAAGCAGCCACAACAAGTATGCAGGGGAGTC
ATTCCAGGAATTTATGATGCTCTGTTTGATATTGAAGCAAAGTGGACCCTTCCAAGGCT
GGGGAGAAGTGAAGAGACAGATTTATGTTGCAGCCTTACAGTGCAGGCAGCTGCAGAGACT
TTGAGTGAAGTAGCTTAAGAGGATTTTATAGAGAATCCGTATTGAATTTGTGTGGTATGTCA
CTCAGAAAGAATCGTAATGGGTATATTGATAAATTTAAAAATGGTATATTTGAAATAAAGT
TGAATATTATATATAA

FIGURE 236

></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA52756

><subunit 1 of 1, 750 aa, 1 stop

><MW: 84305, pI: 6.93, NX(S/T): 10

MWLLHETDSAVATARRPRWLCAGALVLAGGFLLGLFGLGWFIKSSNEATNITPKHNMKAFL
DELKAENIKKFLHNFTQIPHLAGTEQNFQLAKQIQSQWKEFGLDSVELAHYDVLLSYPNKTH
PNYISIINEDGNEIFNTSLFEP PPPGYENVSDIVPPFSAFSPQGMPEGDLVYVNYARTEDFF
KLERDMKINCSGKIVIARYGKVFGRGNKVNAQLAGAKGVILYSDPADYFAPGVKSYPDGWNL
PGGGVQRGNIILNLNGAGDPLTPGYPAN EYAYRRGIAEAVGLPSIPVHPIGYYDAQKLLKMG
GSAPPDSSWRGSLKVPYNVGPFTGNFSTQKVKMHIHSTNEVTRIYNVIGTLRGAVEPDRYV
ILGHRDSWVFGGIDPQSGAAVVHEIVRSFGTLKKEGWRPRRTILFASWDAEEFGLLGSTEW
AEENSRLQLQERGVAYINADSSIEGNYTLRVDCTPLMYSLVHNLTKELKSPDEGFEGKSLYES
WTKKSPSPFEFSGMPRI SKLGS GNDFEVFFQRLGIASGRARYTKNWETNKFSGYPLYHSVYET
YELVEKFYDPMFKYHLTVAQVRGGMVFELANSIVLPFDCRDYAVVLRKYADKIYS ISMKHPQ
EMKTYSVSFDLSLFAVKNFTEIASKF SERLQDFDKSNPIVLRMMNDQIMFLERAFIDPLGLP
DRPFYRHVIYAPSSH NKYAGESFPFGIYDALFDIESKVDPSKAWGEVKRQIYVAAFTVQAAAE
TLSEVA

Signal sequence:

amino acids 1-40

N-glycosylation sites.

amino acids 76-80, 121-125, 140-144, 153-157, 195-199, 336-340,
459-463, 476-480, 638-642

Tyrosine kinase phosphorylation sites.

amino acids 363-372, 605-613, 606-613, 617-626

N-myristoylation sites.

amino acids 85-91, 168-174, 252-258, 256-262, 282-288, 335-341,
360-366, 427-433, 529-535, 707-713